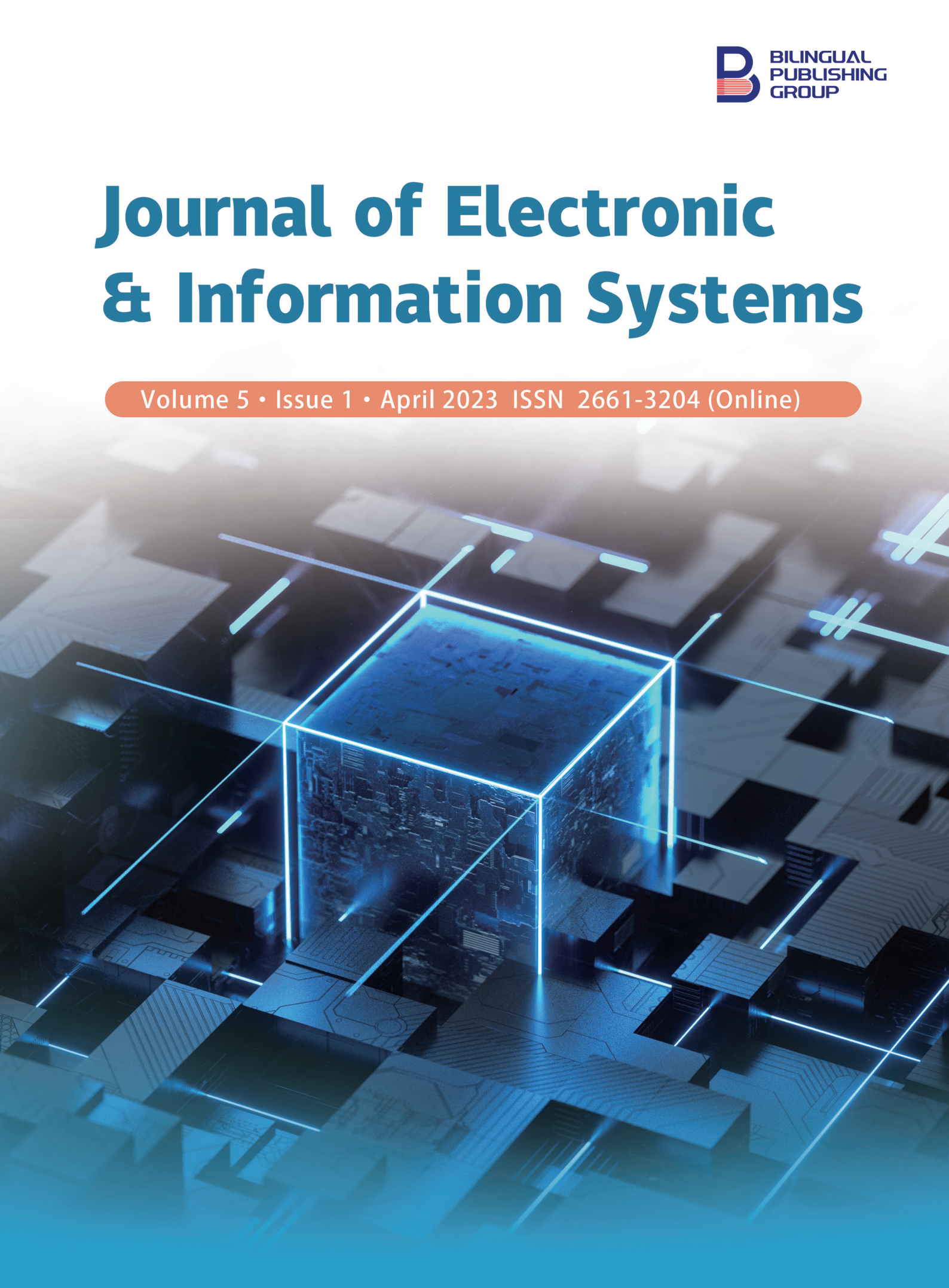
BILINGUAL PUBLISHING GROUP

Journal of Electronic & Information Systems

Volume 5 • Issue 1 • April 2023 ISSN 2661-3204 (Online)



Volume 5 Issue 1 • April 2023 • ISSN 2661-3204 (Online)

Journal of Electronic & Information Systems

BILINGUAL PUBLISHING GROUP

BILINGUAL PUBLISHING GROUP

Volume 5 | Issue 1 | April 2023 | Page1-66

Journal of Electronic & Information Systems

Contents

Articles

Short Communication

BILINGUAL PUBLISHING GROUP

Journal of Electronic & Information Systems
https://journals.bilpubgroup.com/index.php/jeis

ARTICLE

Design of Multi-gas Monitoring Device for Indoor Air Quality

***Serri Abdul Razzaq Saleh*[1], *Huda Jamal Jumaah*[1*], *Zainab Ali Khalaf*[2], *Sarah Jamal Jumaah*[3]**

[1] *Department of Environment and Pollution Engineering, Technical Engineering College-Kirkuk, Northern Technical University, Kirkuk, 36001, Iraq*

[2] *College of Electronic Engineering, Ninevah University, Mosul, 41002, Iraq*

[3] *Center of Disabled Rehabilitation, Iraqi Ministry of Health, Kirkuk, 36001, Iraq*

ABSTRACT

Besides the need for low-cost instruments for air pollution measurement and detection, nowadays there are many concerns about air pollution due to the fast changes and used technologies. This research was applied using an MQ_2 gas detector, and microcontroller/Arduino-Uno. The design steps included bonding and connecting readymade sensors, coding, and finally testing the device. Testing has been conducted in Environment and Pollution Engineering Department laboratories, at the Technical Engineering College of Kirkuk. This study proposed the use of an MQ_2 sensor for multi-gas rate detection which can exist indoors. The system uses also a DHT22 sensor for measuring environment temperature and humidity. The sensors are connected to Arduino and LCD to present data on LCD by powering the system with external power. Overall, the testing was conducted, and the device served as a measuring tool for indoor air as an accurate multi-gas rate detector.

Keywords: MQ_2 sensor; Arduino Uno; Indoor air quality (IAQ); Multi gas; DHT22

1. Introduction

While official attempts to control air contamination have usually concerned the outdoor environment, it is now apparent that raised pollutant levels are more common inside public buildings and houses [1]. Indoor pollution has been classified among the highest five environmental hazards to public health [2]. People are spending their time in homes, offices, or any workplace in long-termly, so it is necessary to

*CORRESPONDING AUTHOR:
Huda Jamal Jumaah, Department of Environment and Pollution Engineering, Technical Engineering College-Kirkuk, Northern Technical University, Kirkuk, 36001, Iraq; Email: huda80@ntu.edu.iq

ARTICLE INFO
Received: 6 January 2023 | Revised: 10 February 2023 | Accepted: 15 February 2023 | Published Online: 24 February 2023
DOI: https://doi.org/10.30564/jeis.v5i1.5390

CITATION
Saleh, S.A.R., Jumaah, H.J., Khalaf, Z.A., et al., 2023.Design of Multi-gas Monitoring Device for Indoor Air Quality. Journal of Electronic & Information Systems. 5(1): 1-9. DOI: https://doi.org/10.30564/jeis.v5i1.5390

monitor environmental air quality which may affect human health [1]. It stances a serious health threat to humans [3-5]. Particularly for children and, persons who spend much time indoors [6]. Besides, the various health problems faced by workers in modern closed office buildings [7], where, levels of some pollutants are greater indoors than outdoor levels, and the personal exposures cannot be characterized sufficiently by outdoor measurements (for some contaminants) [1].

Also, the dependency on solid fuels and incompetent stoves has other, far-reaching concerns of health, environment, and economic development [6]. So, presented studies specify that Indoor Air Pollution (IAP) produced from cooking and heating causes notable health impacts in addition to penetrated outdoor pollution [8]. Therefore, a comprehensive strategy for the detection of indoor pollution, health impacts, control methods, and policy alternatives must be provided [1].

To evaluate the air quality it required an air pollutant concentration is introduced from a model or a monitoring tool [9]. In air quality monitoring aspects, the commercial availability of micro-sensors technology is allowing the adoption of low-cost sensors [10-12]. Indoor Air Quality (IAQ) is considered an important feature to be monitored and controlled for health aspects and comfort [2]. Various conventional devices are presented to monitor contaminants to detect indoor air quality, most of these tools are expensive and inadequate in providing high-resolution spatial-temporal measurements, which are necessary to identify the peak exposure concentrations and detect the sources responsible for the indoor pollution [4]. Moreover, risky gases like methane and propane are combustible so cause explosions when restricted in a closed place [13]. Studies on the early detection of fires were applied by using sensors [14]. The sensor MQ_2 detects gas leakages in small or large places and is a protective system against explosions and fires [15]. Generally, gas sensors are categorized by their operational system, the most common are thermal, electrochemical, ampere-metric, potentiometric, and optical sensors [16]. MQ_2 sensors detect harmful gases [17]. Where, it has a small heater along with an electrochemical detector that reacts with a range of gases, It's used to detect some toxin gases and radiations in case of any poisonous gases or radiation in industrial and living areas [18]. In spite of their limitations, the MQ sensor is used in many applications and Arduino-based projects. For instance, based on air quality we can turn a fan off, or make an alarm system to warn us about smoke [15]. Typically, the Arduino-based project needs slight information about programming or little knowledge of electronics theories and code usage. This may involve indoor and outdoor air quality monitoring using wireless applications [19]. According to the US Environmental Protection Agency, people's exposure to indoor pollutants can on occasion be higher than outdoor pollutants exposure due to interior buildings accumulating and concentrating contaminants that result from furnishings, and the activity of these building's occupants. In effect, indoor pollutants were considered a serious environmental risk to human health and most importantly, if we take into account that most public currently spends higher than 80% of their time in an artificial environment [2].

Indoor air quality is subject to pollution with chemicals, or any toxins that impact human health, as internal air contaminants are related to numerous health risks, such as asthma. In recent years it has been confirmed that interior air quality indoors is rather contaminated than in open air or outdoors. Actually, air types indoors are rather hazardous than outdoor air types. Quality of indoor air methods analysis involves air sampling, building surface samples, and indoor airflow modeling. The process of analyzing indoor air helps in understanding the causes of contamination, then can control it or remove undesirable materials from the air. Ideal air-type measuring systems involve expensive instrument usage. The essential purpose of this paper was to presen t an instrument for measuring the indoor quality of air along with room temperature and humidity, where it was tested in a laboratory/building in a college set-up.

2. System design

The project in the main is based on Arduino run

by a power supply provided for the system via USB. The used sensor for detecting multi gases rates is the MQ_2 sensor which will be set up and bonded to the Arduino with humidity and temperature sensor. To sense the existence of any gas, and measure the temperature of ambient and relative humidity a signal will be sent to the Arduino, then based on programming the data will be collected.

The upcoming explanation involves the details of the system and its components.

2.1 Gas sensor MQ_2

Data acquirement related to gaseous contaminants, and air quality is extensive [20]. The portable air pollution sensing instruments enable the determination of exposure to air pollutants [21]. The sensor MQ_2 is used to measure the smoke and combustible gasses [13,19,22-25], such as propane, methane gas, hydrogen gas, alcohol, and carbon monoxide [19,23-27]. It is useful for detecting leakage of gases inside buildings as an indoor air quality detector [19,22]. It is a metal oxide semiconductor kind of sensor (chemiresistor), which contains a voltage divider network that measures released gas concentration with the ability to detect 200-10000 ppm gas concentration range, working on 5 Volt DC [22-24].

Figure 1 represents the used MQ_2 sensor module. Furthermore, the technical specifications of the used MQ_2 sensor module are described in **Table 1**.

Figure 1. The used MQ_2 sensor module [23].

The MQ_2 sensor is circumfluent by two layers of a fine mesh made of stainless steel named an anti-explosion network. Because MQ_2 senses flammable gases, this network will prevent the heater element from causing an explosion inside the sensor. This mesh is bonded to the body of the sensor with a clamping ring plated by copper as shown in **Figure 1**.

Table 1. The technical specifications of the used MQ_2 sensor module [23].

Sensor details	Specifications
Sensor operating voltage	5 Volt
Electrical load resistance	20 Kilo ohm (KΩ)
Heat resistance	33Ω ± 5%
Heat consumption	<800 mw
Sensing resistance	10 KΩ-60 KΩ
Concentration measure range	200-10000 ppm

As well, works by filtering out the suspended particles to ensure only gaseous pollutants pass into the chamber in addition to protecting the sensor. **Figure 2** represents the internal structure of the MQ_2 sensor when an outer part is set aside. It seems like a star-shaped structure consisting of an element of sensing and connecting legs. Beyond the leads, two of them (H) work to heat the sensor and are bonded by a conductive alloy (Nickel-Chromium coil). The other four leads (A) and (B) responsible for the output signal are bonded by platinum wires. The wires are bonded to the element of sensing and transfer little changes in the current that goes through the sensing element. The element of sensing is a tubular shape and manufactured as (Al_2O_3) Aluminum Oxide based ceramic and also coated by Tin Dioxide (SnO_2). This material (SnO_2) is the most sensitive part of flammable gases. However, the ceramic is heating the sensor area constantly at working temperature and is only increasing the heating efficiency.

The sensor provides an analog output voltage that changes proportionally with the smoke or gas concentration. A higher gas concentration results in high output voltage, while a lower gas concentration results in a lesser output voltage [23].

Moreover, **Figure 3** represents the sensor MQ_2 pinout, where VCC supplies power for the module and it could be connected to 5 volt output from the Arduino. GND: The ground pin must be connected to the Arduino GND pin. DO presents a digital form of the existence of flammable gases. AO presents an analog output voltage that is proportionate to the smoke or gas concentration [23,28].

Since the MQ_2 is a heater-driven sensor, the sensor calibration may drift if you leave it for a long time. For the first usage after being left in storage for one month or more, for example, you should fully heat it for one or two days to achieve high accuracy in measuring. While if it recently has been utilized, the sensor only takes five to ten minutes to fully heat. Over the heating time, the MQ_2 sensor typically measures high but gradually becomes low and decreases until measurements are stabilized [23].

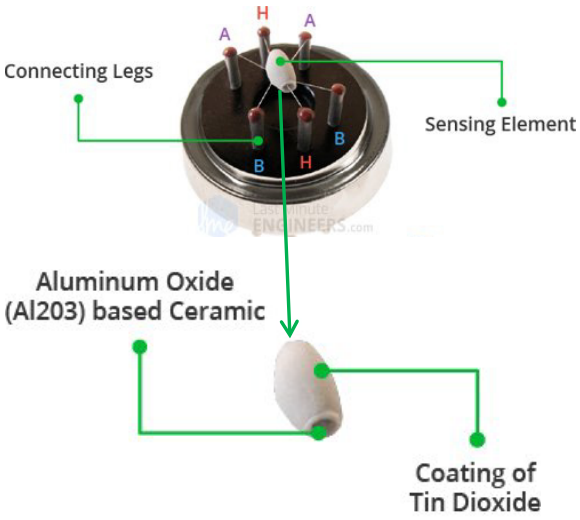


Figure 2. The internal structure of the MQ_2 sensor [23].

Figure 3. The sensor MQ_2 pinout [23].

2.2 Temperature and humidity sensor/DHT22/AM230x/RHT0x

The DHT22 sensor deals with both room temperature and humidity. The occupied high temperature is + 80 °C and the low temperature of –40 °C while humidity ranges from 0-100%. The measuring accuracy is 0.5 °C for temperature, and 2% for humidity accuracy. DHT22 is a digital sensor comprising a temperature measurement called a thermistor and another sensor called a capacitive sensor to determine the humidity. This sensor constituents 4 pins: Pin 1 represents the power pin, pin 2 represents the data pin, pin 3 represents the NULL pine, and pin 4 represents the ground pine. It needs a voltage supply between 3.3-6 volts. This module DHT22 has the best specifications which consider the highest expensive type. Though it has high accuracy and workability at a higher rate of temperature [29]. **Figure 4** represents the DHT22 sensor module. Furthermore, the technical specifications of the DHT22 sensor module are described in **Table 2**.

Figure 4. DHT22 sensor module [29].

Table 2. The technical specifications of the DHT22 sensor module [29].

DHT22 details	DHT22 sensor specifications
Power	3.3 to 6 V
Output signals	Digital signals by single-bus
Sensor element	Polymer capacitor
Range	Humidity (0-100)% and temperature (–40~80) °C
Accuracy	Humidity ±2% and temperature <±0.5 °C
Sensitivity	Humidity 0.1% and Temperature 0.1 °C
Sensing Time	Two seconds
Interchangeability	Fully

Moreover, **Figure 5** represents the DHT22 sensor Pinout, where VCC: Supplies the power to the module, and it is recommended 5 V. Data pin: To communicate the sensor with the microcontroller. NC: Will not be connected. GND: Will be connected to the Arduino ground. Like the DH11 temperature sensor, it is easy to bond the DHT22 sensor to Arduino. It has quite a long 0.1"-pitch pin so it can simply be attached to any board. After powering the sensor by 5 V and connecting ground to ground then, bond the data pin with digital pin #2 [29].

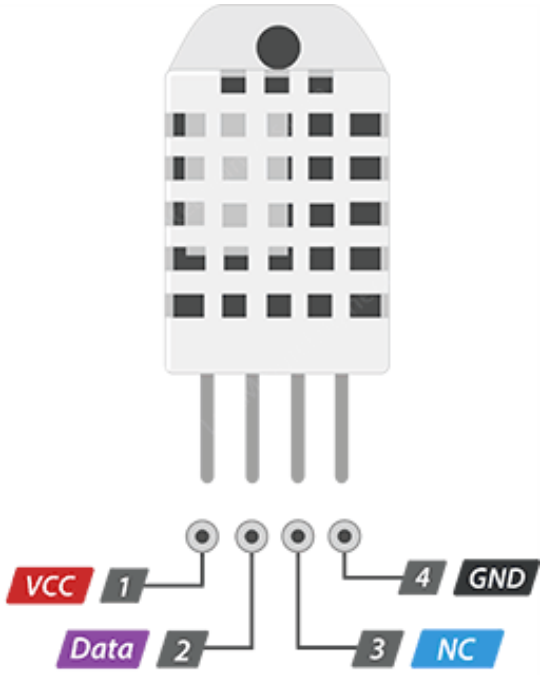


Figure 5. The DHT22 sensor pinout [29].

2.3 Arduino Uno R3

In this study, we presented an air detector device, which consists of a gas sensor based on Arduino. Arduino Uno has the option to acquire info on environs in excess of many sensors then could be practiced simply. The utilized Arduino in our design is Arduino/Uno R3. It comprises the following parts: Microcontroller, USB port, power pins, DC power jack, analog pins, reset bottom, and digital inputs-output pins. **Figure 6** represents Arduino UNO R3.

Arduino Uno considers the preferred board for beginning with electronics, programming, and coding. The Uno is highly robust that can be used mostly. The Arduino Uno is a highly-rated used one and the standard board in the entire Arduino family. The board functional voltage is almost 5, and can be run using a laptop or PC by a USB connector besides an AD-to-DC electric plug or battery operated by an external power supply, here used a battery provides 9 V.

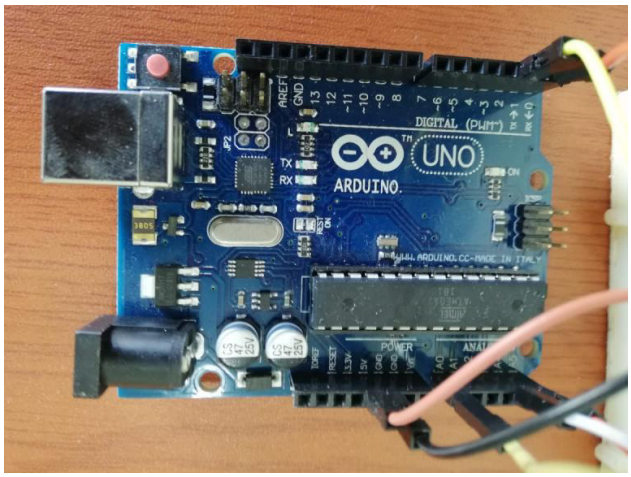


Figure 6. Arduino UNO R3.

2.4 Liquid color displays Lcd-i2C

For any microcontroller, inspecting and combining the characters on Lcd, getting familiar with the data on the Lcd, and configuring it, is the essential task, and among microcontrollers, Arduino is the better one. Arduino is an extraordinary step to interface the LCDs. **Figure 7** represents the Lcd-i2C.

Figure 7. The Lcd-i2C.

2.5 Power supply

In order to run the board a power supply (6-20 V), the power source must be set to turn on the board. The range set is 7-12 V. 9 V functions excellently can be used. Also, Arduino can be organized by a USB link to the computer. Nevertheless, you need to collect outlines that will be self-governing and organized by a battery. It's best for controlling Arduino definitely with a battery instead of voltage controllers which need additional power.

3. Procedure

The study procedure included two steps: Design and testing. To design the presented device for indoor air testing we used Arduino, an MQ_2 sensor, DH22 humidity, and temperature sensor. The representative form of the suggested system is shown in **Figure 8**, which represents the MQ_2 gas detection circuit diagram.

In this design, and based on **Figure 8** the MQ_2 gas detection circuit diagram, the procedure involved connecting the gas sensor MQ_2 and DHT22 temperature and humidity sensor on the Arduino board beside the supplementary apparatuses. The Arduino Uno is programmed using the software of Arduino (IDE) version 1.8.9.

Primarily, each sensor was linked with the Arduino, later we uploaded all codes. After the experiment and the success of the obtained results of independently applied codes other sensors with supplementary components of the system were linked with Arduino. We similarly combined all the codes and applied them with the display type on an Lcd monitor. Consequently, the obtained results will appear as the MQ_2 gas value PPM, temperature value in °C and humidity value in %.

Moreover, **Figure 9** represents the operating Arduino software with uploaded codes.

To call the tools from the Arduino library, first the libraries are defined inside the Arduino. The second step involves defining the sensor type and defining the inputs for the sensors. Here, it starts displaying the temperature and humidity values in fractional form, then it starts displaying the data on the screen and shows the results as numbers. Our designed device will detect air quality based on butane, propane, hydrogen, smoke, methane, alcohol, and carbon monoxide, which will be displayed as a rate of multi-gases values.

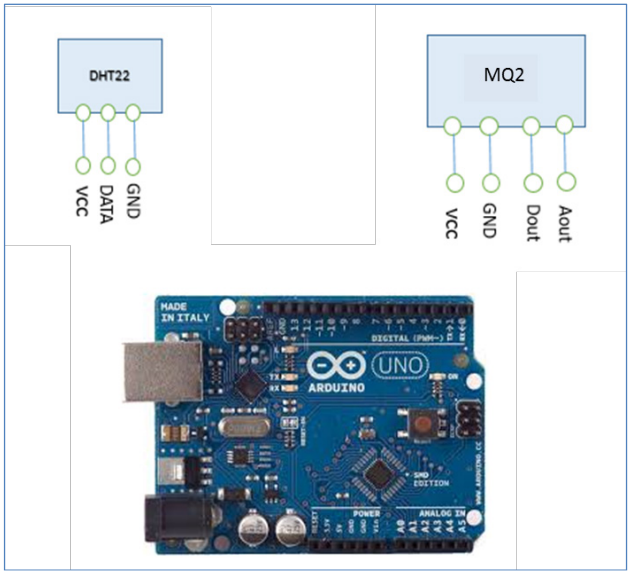


Figure 8. The MQ_2 gas detection circuit diagram.

Figure 9. The operating Arduino software with uploaded codes.

To test the device four testing areas have been selected inside the building. To ensure the workability of the system it was tested at different periods. The measurements were conducted from February to April 2021. The data were collected at four educational laboratories. The range of MQ_2 was within 301 PPM to 807 PPM which was programmed to appear as Smoke on the LCD.

The collected data appeared different readings. The highest values were detected in the Chemistry and Water Pollution laboratories at the Technical Engineering College of Kirkuk.

Maximum humidity was 95% detected in February 2021. While the minimum humidity was 10% detected in April 2021. Moreover, the maximum temperature was 35 °C in April 2021, and the minimum temperature was 16 °C in February 2021.

4. Experimental and practical results

We have been engaged in the design of a portable air quality monitor (MQ_2 gas monitoring) with low-cost and useful, using the Arduino platform.

The used sensors and Arduino are commercially available in the market. Where the device is operated either from the power provided by the computer or from an external supply as a 9-volt battery, and the result appears as a measure of indoor air pollution on the display screen in units Part Per Million PPM which refers to the rate of multi gases values, with the room temperature and rate of relative humidity displayed during the measurement period.

Figure 10 displays the comprehensive experimental progress.

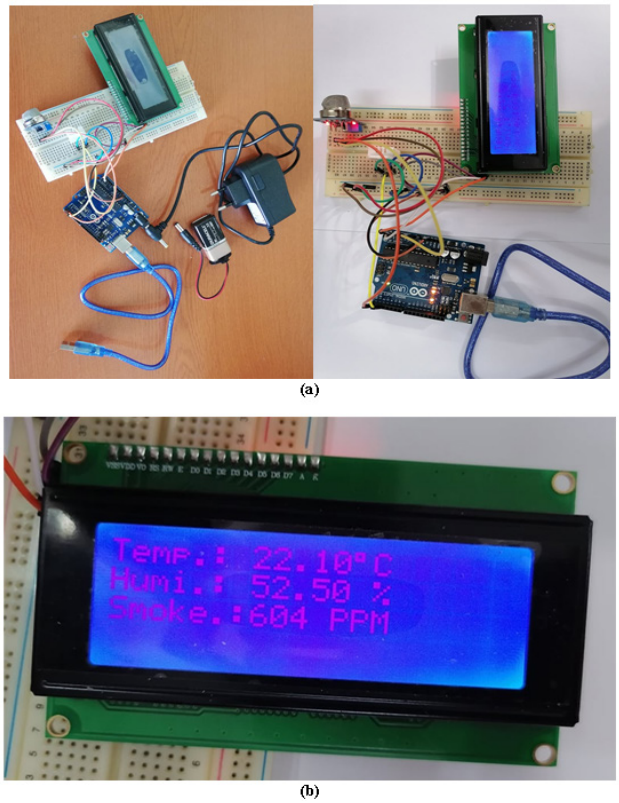


Figure 10. The comprehensive experimental progress: (a) Apparatuses of MQ_2 multi-gas monitoring device, (b) output results.

Where **Figure 10(a)** represents apparatuses of the MQ_2 multi-gas monitoring system, involving all components required for starting the sampling of indoor air. While **Figure 10(b)** represents output results, which it displayed on Lcd as Smoke, value in PPM for MQ_2 gas as a multi-gas rate, Temp. for temperature value in °C and Humi. for humidity value in %.

The data collected by the device were compared with other data for the purpose of verifying the accuracy of the readings. The results were identical and somewhat equal to the readings and measurements.

The major advantages of this indoor monitoring device can be described as a simple design consisting of readymade apparatuses the calibration can be made easily, obtainable sensors, and open source codes, the device can be extended by programming additional sensors, besides its lightweight, it is portable and it easily samples the data, and design cost is low in comparison to other instruments.

Nowadays, getting indoor air quality and pollutants data by means of Arduino-based projects has become popular and useful. Sensor gathering for air quality measurements demonstrates the prospect and potential for many monitoring purposes and sensing [11].

5. Conclusions

In this presented paper, a monitoring instrument for IAQ monitoring applications was designed. The system is established utilizing a low-cost gas sensor MQ_2 that is commercially available in the market in addition to temperature and humidity sensors. Relying on the open-source microcontroller development platform using Arduino the device has been developed. The measuring ability of the system was achieved by testing in different locations and periods. Such a device is very useful in monitoring air quality conditions inside buildings to better understand the current status of air quality as well as to study the long-term impacts of bad and polluted air on public health. Furthermore, the developed device based on the MQ_2 sensor has potential applications in many aspects that required the detection of some harmful gas levels in the air.

Conflict of Interest

There is no conflict of interest.

Acknowledgment

The authors thank the Engineering Academic Office and Dr. Omer Alazzawi for providing the sensors that were used in the project. The authors also thank the laboratory officials in the Environmental Engineering Department at the Technical College of Engineering in Kirkuk for their assistance in taking air samples inside the laboratories during the study period and providing the supporting equipment to evaluate the measurements.

References

[1] Spengler, J.D., Sexton, K., 1983. Indoor air pollution: A public health perspective. Science. 221(4605), 9-17. doi: https://www.science.org/doi/abs/10.1126/science.6857273.

[2] Pitarma, R., Marques, G., Ferreira, B.R., 2017. Monitoring indoor air quality for enhanced occupational health. Journal of Medical Systems. 41(2), 1-8. doi: 10.1007/s10916-016-0667-2.

[3] Abraham, S., Li, X., 2014. A cost-effective wireless sensor network system for indoor air quality monitoring applications. Procedia Computer Science. 34, 165-171. doi: 10.1016/j.procs.2014.07.090.

[4] Kumar, P., Skouloudis, A.N., Bell, M., et al., 2016. Real-time sensors for indoor air monitoring and challenges ahead in deploying them to urban buildings. Science of the Total Environment. 560, 150-159. doi: 10.1016/j.scitotenv.2016.04.032.

[5] Brilli, F., Fares, S., Ghirardo, A., et al., 2018. Plants for sustainable improvement of indoor air quality. Trends in Plant Science. 23(6), 507-512. doi: 10.1016/j.tplants.2018.03.004.

[6] Bruce, N., Perez-Padilla, R., Albalak, R., 2000. Indoor air pollution in developing countries: A major environmental and public health challenge. Bulletin of the World Health Organization. 78, 1078-1092.

[7] Samet, J.M., Marbury, M.C., Spengler, J.D., 1987. Health effects and sources of indoor air pollution. Part I. American Review of Respiratory Disease. 136(6), 1486-1508. doi: 10.1164/ajrccm/136.6.1486.

[8] Smith, K.R., 2002. Indoor air pollution in developing countries: Recommendations for research. Indoor Air. 12(3), 198-207.

[9] Jumaah, H.J., Ameen, M.H., Kalantar, B., et al., 2019. Air quality index prediction using IDW geostatistical technique and OLS-based GIS technique in Kuala Lumpur, Malaysia. Geomatics, Natural Hazards and Risk. 10(1), 200-207. doi: 10.1080/19475705.2019.1683084.

[10] Piedrahita, R., Xiang, Y., Masson, N., et al., 2014. The next generation of low-cost personal air quality sensors for quantitative exposure monitoring. Atmospheric Measurement Techniques. 7(10), 3325-3336.

[11] Jumaah, H.J., Kalantar, B., Halin, A.A., et al., 2021. Development of UAV-based PM 2.5 monitoring system. Drones. 5(3), 60. doi: 10.3390/drones5030060.

[12] Karagulian, F., Barbiere, M., Kotsev, A., et al., 2019. Review of the performance of low-cost sensors for air quality monitoring. Atmosphere. 10(9), 506. doi: 10.3390/atmos10090506.

[13] Jualayba, M., Regio, K., Quiozon, H. (editors), et al., 2018. Hazardous gas detection and notification system. Proceedings of IEEE 10th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM); 2018 Nov 29—Dec 2; Baguio City, Philippines. USA: IEEE. p. 1-4. doi: 10.1109/HNICEM.2018.8666358.

[14] Kamelia, L., Ismail, N., Firmansyah, A.A., 2019. Fire disaster early detection system in residential areas. Journal of Physics: Conference Series. 1402(4). doi: 10.1088/1742-6596/1402/4/044001.

[15] Bogdan, M., 2018. Gas detector using Arduino and LabVIEW. Proceedings of the 13th International Conference on Virtual Learning ICVL.

315-318. Available from: http://c3.icvl.eu/papers2018/icvl/documente/pdf/section2/section2_paper46.pdf.

[16] Villa, T.F., Salimi, F., Morton, K., et al., 2016. Development and validation of a UAV based system for air pollution measurements. Sensors. 16(12), 2202.

[17] Waworundeng, J.M.S., 2018. Prototype of gas detector with IoT platform for notification and monitoring system. Abstract Proceedings International Scholars Conference. 6(1), 160. doi: 10.35974/isc.v6i1.1243.

[18] Srinivas, C., Ch, M.K., 2017. Toxic gas detection and monitoring utilizing Internet of things. International Journal of Civil Engineering and Technology. 8(12), 614-622.

[19] Heyasa, B.B.L., Van Ryan Kristopher, R.G., 2017. Initial development and testing of microcontroller-MQ_2 Gas sensor for university air quality monitoring. IOSR Journal of Electrical and Electronics Engineering. 12(3), 47-53. doi: 10.9790/1676-1203024753.

[20] Villa, T.F., Gonzalez, F., Miljievic, B., et al., 2016. An overview of small unmanned aerial vehicles for air quality measurements: Present applications and future prospectives. Sensors. 16(7), 1072.

[21] Maag, B., Zhou, Z., Thiele, L., 2018. W-air: Enabling personal air pollution monitoring on wearables. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies. 2(1), 1-25.

[22] Thamaraiselvi, D., 2020. Envionment monitoring system using IoT. Mukt Shabd Journal. 9(5), 915-920.

[23] Last Minute Engineers. How MQ_2 Gas/Smoke Sensor Works? & Interface it with Arduino [Internet] [cited 2021 Mar 7]. Available from: https://lastminuteengineers.com/mq2-gas-senser-arduino-tutorial/.

[24] ELPROCUS. MQ_2 Gas Sensor—Working Principle & Its Applications [Internet] [cited 2021 Mar 7]. Available from: https://www.elprocus.com/an-introduction-to-mq2-gas-sensor/.

[25] Krishnamoorthy, R., Krishnan, K., Bharatiraja, C., 2021. Deployment of IoT for smart home application and embedded real-time control system. Materials Today: Proceedings. 45, 2777-2783. doi: 10.1016/j.matpr.2020.11.741.

[26] Jayakumar, D., Ezhilmaran, R., Balaji, S., et al., 2021. Mobile based gas leakage monitoring using IOT. Journal of Physics: Conference Series. 1717(1), 12068.

[27] Devi, K.I., Meivel, S., Kumar, K.R., et al., 2021. A survey report of air polluting data through cloud IoT sensors. Materials Today: Proceedings. Available from: https://linkinghub.elsevier.com/retrieve/pii/S2214785320403384.

[28] Setiawan, F.N., Kustiawan, I., 2018. IOT based air quality monitoring. IOP Conference Series: Materials Science and Engineering. 384(1), 12008.

[29] Last Minute Engineers. Interface Itwith Arduino [Internet] [cited 2021 Mar 7]. Available from: https://lastminuteengineers.com/dht11-dht22-arduino-tutorial/.

BILINGUAL PUBLISHING GROUP

Journal of Electronic & Information Systems
https://journals.bilpubgroup.com/index.php/jeis

ARTICLE

The Theory on Thing's Limits. Part 2: A Brief Analysis of the New Knowledge of Newton's First Law

Jian Ding

Integrated Electronic Systems Lab Co. Ltd., Jinan, Shandong, 250100, China

ABSTRACT

According to the norm of identifying truth in this theory, and Newton's first law as a basis that can look at the overall situation, and by virtue of the electron storage ring as an experimental fact, it is pointed out: Only in reality can there be inertia. Inertia represents the continuity of the development of things. As the speed gradually approaches the *c*, the particle's mass also approaches zero along with its static mass due to the impact of electromagnetic radiation, which is exactly the root where the energy shrinkage effect of high-speed particles comes from, and also the primary factor causing the spectrum redshift. Therefore, the Big Bang theory is wrong. All photons are produced from high-density particles through electromagnetic radiation. Wherever there is fluctuation, there must be mass, and vice versa. This is the correct understanding of "wave-particle duality". No matter whether the high-speed electrons or the photons produced by them all have different static masses, their charge-mass ratio is always the same physical constant, and not affected by relativistic effects and electromagnetic radiations. This is the true internal mechanism to constitute the uncertainty principle, and conforms to the experimental facts related to it. It can be proved that in a constant magnetic field, the high-speed electron or photon having a relatively large curvature radius has a high moving speed and less mass, energy and wave frequency. Since Einstein used the absolute space-time established by Newton as the criterion and came to the conclusion that the relative space-time was curved, then he should no longer make circular arguments, that was, used the relative space-time as the criterion, to change the unit length and time established by the absolute space-time.

Keywords: Mechanics; Relativity; Electron; Photon; Electromagnetic radiation; Energy shrinkage

*CORRESPONDING AUTHOR:
Jian Ding, Integrated Electronic Systems Lab Co. Ltd., Jinan, Shandong, 250100, China; Email: jiandus@163.com

ARTICLE INFO
Received: 16 March 2023 | Revised: 23 March 2023 | Accepted: 28 March 2023 | Published Online: 6 April 2023
DOI: https://doi.org/10.30564/jeis.v5i1.5548

CITATION
Ding, J., 2023. The Theory on Thing's Limits. Part 2: A Brief Analysis of the New Knowledge of Newton's First Law. Journal of Electronic & Information Systems. 5(1): 10-19. DOI: https://doi.org/10.30564/jeis.v5i1.5548

1. Introduction

This article is the second part of the full text of "The theory of thing's limits". When quoting from the content of the full text's Part 1 [1], use "P1" to index. The rest may be inferred by analogy. In P1, the characteristic of truth has been explained, that was, it could not be proved by empirical methods, and only be approached gradually by repeated practice. And according to the principle of inertia, the norm for identifying truth has been established. Based on this, we can break through the bondage of finite thinking, from the quantitative change of real space have gone deep into the qualitative change of the ideal realm, and extend the philosophy of materialism to the category of metaphysics. Not only have restored the true nature of metaphysics, but also achieved the goal that Sir Isaac Newton pursued throughout his whole life. That was, theological (metaphysical) thought and scientific ideal were closely related, which was an organic whole, and the wisdom in one domain might enlighten the wisdom in another [2].

This means that metaphysics has neither divorced from practices, nor just observed objective things with a one-sided, isolated and static way of thought. In people's daily life, everyone has used it, just not deliberately reflected on it.

The contents contained in metaphysics, such as noumenon, axiom or postulation, absolute motion, etc., all can be called truths. They must have absoluteness and immutability, and do not exist in reality. So there is no direct connection between them, but they have the continuity with objective things in reality. That is to say, the differences between all truths and objective things can reach an arbitrarily small value.

This means that once a certain truth is mastered, it is equivalent to standing in a position that can look at the overall situation. Thereby getting rid of the dilemma of placing oneself in the midst of events, which is unable to look at the overall situation clearly, so difficult to grasp the correct direction of research and always using a one-sided view to treat the overall problem. At this time, again looking back at the objective things in reality, that is just like Sir Newton's famous saying: "If I have seen further than others, it is by standing upon the shoulders of giants."

As the introduction of the concept of absolute motion, Newton's first law of motion can be called the starting point of classical physics, which has been widely applied to physical space in reality, and is the best example. In other words, if compared with any one of the "giants" who are only confined to the real physical space and determined to "gives no basis for the introduction of the concept of absolute motion" [3], he would be as always standing a little higher and seeing a little farther. Only in this way, it can be possible to realize the goal of mutual enlightenment between natural science and metaphysics [4].

2. The truth contained in Newton's first law

Newton's first law of motion is also known as the law of inertia. It is scientifically to clarify the two physical concepts of force and inertia, and correctly explain the relationship between the force and the state of motion, as well as to put forward that all bodies have the property to keep their state of motion unchanged, that is inertia. It is expressed as [5]: Everybody perseveres in its state of rest, or of uniform motion in a right line, unless it is compelled to change that state by forces impressed thereon.

Obviously, the body described therein is unaffected by forces, no matter whether it is stationary or moving along a straight line in uniform motion, its motion state is absolute (or ideal), does not exist in reality, and belongs to the category of metaphysics. And in reality, no matter what kind of physical problems you study, the law as a shadow always follows. Does it seem to prompt us to reflect on what the physical phenomena, in reality, are through what kind of a way for the introduction of the concept of absolute motion provides a basis?

In this regard, the ideas and methods for the formation of "The theory on thing's limits", in P1, as well as the norm of identifying truth, in which discussions have already given the answer. That is to say, Newton's first law has also told us that only

in real space can inertia exist. Inertia represents the continuity of the development of things. It can be inferred from this that as long as in reality, between any two things a direct or indirect causality is bound to find through continuity. Next, use again the above-mentioned ideas, methods and norms to give a further discussion of several concepts involved in this law.

First of all, the "right line" defined in the law, which is an absolutely straight line, does not exist in reality and belongs to the category of metaphysics. There is an error between any straight line in reality and this absolute straight line. But there must be a straight line in reality, and the error between it and this absolute straight line can be an arbitrarily small value. Obviously, using this straight line in reality to standardize space-time is the most accurate, but there is still an error after all. Therefore, the main use of this straight line is to abstract Newton's that absolute error-free straight line.

So by the definition, the straight-line distance S is equal to the velocity V of uniform linear motion times the time spent T (S = VT). The coordinate system thus established describes the absolute time and absolute space of Newtonian mechanics, which are referred to as absolute space-time for short. Only with this as the criterion, can we start from the consensus of absolute no error to understand the changes of objective things in reality.

Einstein since used absolute space-time as the criterion and came to the conclusion that the relative space-time (i.e., the space-time in reality) was curved, he should no longer use the relative space-time as the criterion to change the unit length and unit time established by the absolute space-time. This kind of circular argument confused right and wrong, which was a logical fallacy, and even would deduce to turn back in time, so would not muddle through certainly in the face of truth [3].

So why is the relative space-time curved? It is because of the existence of matter in there. With matter, there must be the interactions of forces among them. According to Newton's first law, due to inertia, the matter's linear motion has been bent or fluctuated, and the velocity also becomes uneven. So using Newton's absolute space-time as the criterion, you will feel that the relative space-time is curved, or the density uneven. And for this, various concepts such as mass, density, inertia, temperature, pressure, speed, frequency, wavelength, momentum, energy, etc., have been defined.

That is to say, as a truth, Newton's first law has told us that when bidirectional reasoning is carried out between it and objective things, these concepts constitute our macro-cognitions of the different states in the reasoning processes. By virtue of the continuity of the objective things associated with them, there are causalities among these macro-cognitions, so that we can understand a certain objective thing more comprehensively, and would not produce the wrong cognitions that use one-sided view to treat the overall problem, such as "Blind Men and the Elephant" [6].

3. A particular case of Newton's first law

This particular case has been mentioned in P1, which is the value *c* of light speed in vacuum. In the absence of any external force, the photon always keeps moving along a straight line at a uniform velocity of *c* = 299792458 (m/sec). This is the limit value of the speed in the universe, which is just higher an arbitrarily small value than the highest light speed in reality. This is an absolute motion and also an absolute frame of reference. A coordinate system which is admitted in mechanics is called as an "inertial system". Based on this, all inertial systems can be defined in reality, including the so-called stationary inertial systems. Just like the absolute point positions of natural numbers on the number axis in P1, the abstractions of these inertial systems do not exist in reality, and belong to the absolute space-time view established by Newton.

Therefore, in this particular case and therein the photon in uniform linear motion does not exist in reality. Such a photon is defined only to show that it has continuity with the photon whose speed is closest to the c in reality. From this, it can be seen that

Newton's first law is called the law of inertia, which precisely emphasizes that objects, in reality, have inertia. Where there is inertia, there must be matter and mass, and vice versa. But in view of the fact that truth must have absoluteness and immutability, the noumena of various concepts belong to the metaphysical category, such as mentioned in this article, of inertia, mass, speed, fluctuation, momentum and energy, etc., among their noumena there are no continuity and no direct relationship. They can only be related through objective things that have continuity with themselves in reality.

So, according to the norm for identifying truth in P1, what else is the difference between the photon that moves at the value c of light speed in a vacuum and a photon that moves at the highest light speed in reality?

As far as the former is concerned, which is the limit value that is starting from objective facts to be confirmed through the process of gradual approaching, does not exist in reality. Hence the former has neither mass ($m = 0$) nor energy ($E = 0$). And the latter is a photon in reality, of course, it has both mass ($m > 0$) and energy ($E > 0$). Otherwise, it means that there is no object of receiving force, so the characteristics of inertia certainly can not be shown.

Since in reality, the latter is bound to be affected by external forces. According to Newton's first law, linear motion has become fluctuations, and the velocity is also not uniform. This is precisely the result of the interaction between the external force and the inertial force of the latter. As a result, the concept of "speed", which has been the velocity of moving along the direction of wavelength, also comes into being with it. Obviously, this is exactly the objective reality we see. But more importantly, Newton's first law has told us that wherever there are fluctuations in reality, there must be matter and mass, and vice versa.

Furthermore, in view of the former that has no mass ($m = 0$), but has continuity with the latter, it can be realized that the mass of the latter is already very small. Here, Newton's first law seems to have given instruction on how to grasp the correct direction of research.

That is to say, as the speed of moving the photon in reality gradually approaches the c, its mass will gradually approach zero, and the moving trajectory approach the (absolutely) straight line, and the speed and velocity also tend to unity. This means that in reality, a photon with a relatively high moving speed should have less mass, energy, and wave frequency than another photon with a relatively low moving speed. Obviously, this is very consistent with the objective facts we have seen, such as the relationship between the energy and wave frequency of visible light.

In this regard, as early as the 1920s, the famous de Broglie matter wave relation [7] has shown us that the momentum of a high-density particle (including electron or photon),

$$\boldsymbol{p} = m\boldsymbol{v} = \frac{h}{\lambda} \tag{1}$$

and its energy,

$$E = m\boldsymbol{v}^2 = \frac{h\boldsymbol{v}}{\lambda} = hf \tag{2}$$

were related to its wavelength λ or frequency f, but could not grow without limit, and must be restricted by Planck's constant h. Among them, λ is the moving speed along the direction of wavelength and $\boldsymbol{v} = \lambda f$. As the cornerstone of the expansion of Newtonian mechanics to wave mechanics or quantum mechanics, the matter-wave relation can be through the continuity with objective things in reality, to rely on each other with Newton's three laws of motion and the value c, and produce causality.

4. Photons are the products of electromagnetic radiation

It is indisputable that a photon in reality has mass ($m > 0$). If there is no mass ($m = 0$), it means that there is no matter. That is to say, if there is no object of receiving force, of course, there is no inertia. As a result, those physical phenomena that should be related to it in classical mechanics, such as inertia, fluctuation, temperature, pressure, momentum and energy, etc., would not appear. Because the interac-

tions among matters are the necessary condition for us to perceive the above-mentioned physical phenomena. On the contrary, as long as you find anything related to these physical phenomena, it must be related to mass.

As far as the photon is concerned, it is a high-speed particle that has been produced by electromagnetic radiation from a high-density particle, such as an electron, proton, neutron, or atomic nucleus, etc. Einstein defined it as the quantum of light, and later it has been called the photon for short. According to the kinetic energy formula of his special relativity [8],

$$E_k = E - E_0 = m_0 c^2 (\gamma - 1) \tag{3}$$

the word "quantum" contains indivisibility between matter and energy, while the word "light" is focused on electromagnetic radiation. That is to say, all high-speed particles that are produced from high-density particles through electromagnetic radiation, should be belonged to the category of photons.

The formula (3) also contains the mass-energy formula (4) and mass-speed formula (5) of his special relativity.

$$E = mc^2 \text{ or } \Delta E = \Delta m c^2 \tag{4}$$

$$m = m_0 \gamma \tag{5}$$

Among them, $E_0 = m_0 c^2$ is the static energy and E the total energy. Both are aimed at the same moving object or particle. And the mass-speed formula (5) reveals that the concept of mass is divided into two parts along with the moving speed v of objects (including various high-density particles) changes. In the above formulas, m is the mass to represent inertia, m_0 is the static mass to represent the quantity of matter, $\gamma = 1/\sqrt{1 - v^2/c^2}$ is the expansion factor, and c is the value of light speed in a vacuum.

As far as photons are concerned, since you can perceive their fluctuations and energy, so according to the previous analysis of Newton's first law, and the formulas (1) and (2), should know that they have inertia and mass. This is the correct understanding of "wave-particle duality". If a consensus can be reached on this, will people be the sake of whether light exactly is a wave or a particle to be a state of endless debate and each air his own views? As for photons that have no mass but energy, such a paradox that confuses the two domains of definition, no one will be to believe it anymore.

In fact, the problem is not so simple. Take random electromagnetic radiation as an example. According to formulas (4) and (5), the energy that every one of radiated photons has will take away a part of the static mass (that is, a certain amount of matter). In this regard, there has been no consensus. And this is exactly the objective fact that modern physics must face, any so-called authoritative statements that have attempted to circumvent it vaguely, will certainly come back here again, and reconfirm. Otherwise, modern physics will be fettered here and difficult to make great progress.

In 1927, German physicist Werner Karl Heisenberg first pointed out that the more precisely the position of some particle was determined, the less precisely its momentum could be known, and vice versa. The principle can also be expressed as: When the momentum $\boldsymbol{p}$ is determined, the speed $\boldsymbol{v}$ (or position) cannot be determined; when the speed $\boldsymbol{v}$ (or position) is determined, the momentum $\boldsymbol{p}$ cannot be determined.

In Heisenberg's time, physicists generally believed that an electron would not be further broken down. For a low-speed moving in a non-relativistic state, the charge, (static) mass, and the charge-mass ratio of each electron all were considered to be physical constants, which was a consensus that people formed one hundred years ago. And now we must realize that the charge and static mass of each electron have been no longer constants due to the impact of electromagnetic radiation.

For the electrons with different static masses, their charge-mass ratio e/m_0 is always the same physical constant. When they are moving at different speeds, the momentums or energies can be the same. So the phenomenon presented is "When the momentum $\boldsymbol{p}$ is determined, the speed $\boldsymbol{v}$ (or position) cannot be determined." Conversely, when they are moving at the same speed, the momentums or energies can be various. So the phenomenon presented is "When the speed $\boldsymbol{v}$ (or position) is determined, the momen-

tum $\boldsymbol{p}$ cannot be determined." This is precisely the internal mechanism of the uncertainty principle. In other words, the uncertainty principle and a large number of experimental data related to it have proven that the charge and static mass of electrons both are not constants [9].

5. The energy shrinkage effect of high-speed particles

Now we must realize that if continue to insist on the static mass of every electron as a constant, the continuities among them will be lost. Only by abandoning this wrong view can we ensure that the uncertainty principle will not offend Newton's first law. That is to say, regarding whether an electron will be broken down further, it should be verified carefully.

In reality, along with the moving speed v of a high-speed particle (electron or photon) gradually approaching the c, the same particle exhibits two effects: On the one hand, according to the formula (5) its mass m keeps getting bigger; and on the other hand, its static mass m_0 keeps getting smaller due to the loss caused by electromagnetic radiation. The result of the final synthesis shows that its mass gradually approaches zero along with its static mass. This is the energy shrinkage effect of high-speed particles [10].

For example, the electron storage rings widely in use at present, are large-scale scientific facilities for studying high-energy physics [11]. In the storage ring, an electron momentum [12],

$$\boldsymbol{p} = m\boldsymbol{v} = e\boldsymbol{B}R \tag{6}$$

can be obtained directly from the equation of the centripetal force $m\boldsymbol{v}^2 / R$ and Lorentz force $e\boldsymbol{B}\boldsymbol{v}$. In formula (6), e is the charge of an electron, $\boldsymbol{B}$ is the magnetic induction intensity, and R is the curvature radius of the moving electron in the ring.

For high-speed electrons, the relativistic effects should be considered. By formulas (5) and (6), the relative speed (7) of an electron can be deduced. In view of the fact that the magnetic induction intensity $\boldsymbol{B}$ and curvature radius R in the storage ring have been designed to the given values, and the electron charge-mass ratio e/m_0 is a constant, it can be calculated by formula (7) that the speed $\boldsymbol{v}$ of every moving electron therein is also a given value, and has nothing to do with its mass.

$$\beta = \frac{v}{c} = \sqrt{\frac{(\frac{e}{m_0})^2}{(\frac{e}{m_0})^2 + (\frac{c}{\boldsymbol{B}R})^2}} \tag{7}$$

As a result, these high-speed electrons with different static masses can move along the circular orbit to do cyclotron motion at a given speed v in the tubular vacuum chamber of the storage ring. Under normal circumstances, the electron beam lifetime in the storage ring can even last for more than a dozen hours. Once this period of time has passed, the electron beam intensity will quickly decay to zero. The reason is that during this period of time most of the static mass $\sum \Delta m_0$ belonged originally to each electron itself has been gradually lost as electromagnetic radiation turned into photons. At this time, the static mass m_0 of each electron in the storage ring has become very small. As a result, when a photon radiated, its impulse has been transferred to the electron and enough pushes the electron to the inner wall of the tubular vacuum chamber of the storage ring [13].

The above example, by virtue of the widely used electron storage ring as an experimental fact, has proved that due to electromagnetic radiation, the charge e of a high-speed electron will follow along with its static mass m_0 to be lost synchronously, and its charge-mass ratio e/m_0 always remains unchanged. This is the energy shrinkage effect of high-speed electrons. The nature of this matter is serious and of great significance, in order to be cautious, which should be verified to facilitate consensus. The behavior of any attempt to circumvent or muddle through, may deduce the wrong conclusion, or cause the loss of research direction. In fact, it is not difficult to verify. As long as you do interference or diffraction experiments on the electron beam in the storage ring before and after this period of time separately, and no need to quantify, just compare the wavelengths of the two qualitatively to distinguish them.

The experimental facts have also revealed that

the electron mass M_0 and elementary charge e_0 in the fundamental physical constants only refer in particular to the statistical values of low-speed electrons when just reach the status where they can leave the atoms. The ratio of the two is the charge-mass ratio e_0 / M_0 in the fundamental physical constants. Due to the existence of energy shrinkage effect of every high-speed electron, in general, its static mass and charge are in the state of synchronous reduction, but the charge-mass ratio (8) has remained unchanged. Therefore, it must be emphasized that the charge-mass ratio e/m_0 of an electron refers to the ratio of its charge to the quantity of matter, which is not affected by relativistic effects and electromagnetic radiation. By the same token, this conclusion also applies to every photon that has been radiated by the electrons in the storage ring.

$$e / m_0 = e_0 / M_0 \tag{8}$$

6. A minimal error or deviation may result in wide divergence

Then, these photons must follow the law of motion determined by formula (7). That is to say, under the premise of the given magnetic induction intensity $\boldsymbol{B}$, the photon with a relatively large curvature radius R must have a relatively high moving speed and relatively less mass, energy and wave frequency. This is consistent with the experiment of decomposing sunlight with a triangular prism, and indicating that the visible light is also composed of charged particles of different sizes. In other words, compared with purple light, the photons of red light have relatively high moving speeds, which should be an objective fact that Newton's first law has tacitly approved.

Therefore, it is emphasized again that the energy shrinkage effect of high-density particles caused by electromagnetic radiation is an objective reality that must be faced in modern physics. Combined with formulas (1) and (2), it can be seen that this is exactly the primary factor causing the spectrum redshift. However, in view of the fact that most mainstream thinking has been still restricted to the cognition that an elementary particle (electron or photon) will not be broken down further, and disturbed by the uncertainty principle, which has not caused enough attention. This primary factor causing the spectrum redshift has actually been replaced by the Doppler Effect, and transformed into that celestial bodies are moving away from us. A minimal deviation may result in a wide divergence. But now, the cosmological redshift seems to have been judged as a correct theory while thinking carefully, a little worry involuntarily... [14].

That is to say, just like the moral in the story "Blind Men and the Elephant", the well-known Big Bang theory is a cognitive error that has been used by a one-sided view to treat the overall problem. The reason for such an error, if traced to the source, is still because it has offended the most fundamental truth in physics, that is, Newton's first law. What on earth has been offended? It is inertia, that is, the continuity of the development of things. Only in real space can inertia exist. In other words, any two things between them in reality, a direct or indirect causality is bound to find through continuity. Otherwise, it would be beyond the scope of real space. It should be noted here that only mathematics has the possibility to run through two different domains of definition, that is, both continuity and discontinuity can exist in that.

Take the uncertainty principle as an example. This is a research conclusion that has been drawn from a large number of experimental facts and reached a consensus. If continue to insist the static mass of every electron is a constant, the price paid is that the continuity among them has been lost, that is, offending Newton's first law. Therefore, such a paradox that celestial bodies were moving away from us, would be deduced. This means that the conclusion drawn can only be the static mass of every electron being different. Yes, it not only conforms to a large number of experimental facts, but also gets the acquiescence of Newton's first law, and everything becomes harmonious.

It was because they did not stand on this basis that could look at the overall situation, so these physicists would have ignored the objective fact around themselves. That was, electromagnetic radiation

could cause an electron to be broken down further. From the two conceptual ambiguities for inertia and charge-to-mass ratio, to the uncertainty principle by muddling through, and even the photon that had no mass but with energy, and so on, various similar "authoritative theories" have wandered in modern physics. And once confronted with the judgement of truth, the behavior of any attempt to circumvent or muddle through, will inevitably become the victim of one's own mistake. Although this process may be long, it is insignificant compared to eternal truth.

Incidentally, as a light source of electromagnetic radiation, the speed of moving electrons in the storage ring, $v > 0.99c$, which is sufficient to prove that the value c of light speed in vacuum is independent of the state of motion of the emitting body. On the contrary, it also proves that photons in reality must have static mass, and their size and moving speed are all different. Because according to Newton's first law, as the speed of moving photon in reality gradually approaches the c, its mass will also gradually approach zero along with its static mass.

7. Conclusions

This article is according to the ideas and methods for the formation of "The theory on thing's limits", in P1, as well as the norm of identifying truth, its characteristic is based on Newton's first law, which is equivalent to standing in a position that can look at the overall situation. Thereby getting rid of the dilemma of placing oneself in the midst of events, which is difficult to grasp the correct direction of research and always using a one-sided view to treat the overall problem. The theory is applicable to all academic categories, and can be combined with objective practices. In terms of testing authoritative theories, clarifying chaos, and deriving new knowledge, etc., there are the following conclusions:

1) Newton's first law is called the law of inertia, which precisely emphasizes that objects in reality have inertia. Where there is inertia, there must be matter and mass, and vice versa. Inertia, represents the continuity of the development of things. Only in reality can there be inertia. In other words, any two things between them in reality, a direct or indirect causality is bound to find through continuity. Otherwise, in the whole reasoning process, there must be a fault beyond the real space, which is to offend Newton's first law.

2) Now that Einstein used the absolute space-time established by Newton as the criterion and came to the conclusion that the relative space-time (i.e., the space-time in reality) was curved, then he should no longer use the relative space-time as the criterion to change the unit length and unit time established by the absolute space-time. This kind of circular argument confused right and wrong, which was a logical fallacy, so would not muddle through certainly.

3) There are matters in space-time in reality, so there must be interactions of forces among them. According to Newton's first law, due to inertia, the linear motion of matter has become fluctuations. Therefore, wherever there are fluctuations, there must be matter and mass, and vice versa. As far as photons are concerned, now that we perceive their fluctuations and energy, should know that they have inertia and mass. This is the correct understanding of "wave-particle duality".

4) The value c of light speed in a vacuum is a particular case of Newton's first law. Based on this, it can be deduced that in reality, a photon with a relatively high moving speed should have less mass, energy, and wave frequency than another photon with a relatively low moving speed. This is consistent with the objective facts we have seen. Because every one of the photons that have been radiated by the electron in the storage ring, has the same charge-mass ratio as an electron, but their static masses are different from one another. Then, these photons must follow the law of motion determined by formula (7). That is to say, under the premise of the given magnetic induction intensity $\boldsymbol{B}$, the photon with a relatively large curvature radius R must have a relatively high moving speed. This is consistent with the experiment of decomposing sunlight with a triangular prism, that is, compared with the purple light, the photons of red light have a relatively high moving speed.

5) The photons in reality, are high-speed particles

produced by high-density particles through electromagnetic radiation. According to formulas (4) and (5), the energy that the radiated photons have will take away a part of the static mass. Therefore, by virtue of the widely used electron storage ring as an experimental fact, it has been proved that due to electromagnetic radiation, the charge e of a high-speed electron would follow along with its static mass m_0 to be lost synchronously, and its charge-mass ratio e/m_0 has remained unchanged. This means that the charge and static mass of electrons have been no longer constants.

6) For the electrons with different static masses, when they are moving at different speeds, the momentums or energies can be the same. So the phenomenon presented is "When the momentum $\boldsymbol{p}$ is determined, the speed $\boldsymbol{v}$ cannot be determined." Conversely, when they are moving at the same speed v, the momentums or energies would be various. So the phenomenon presented is "When the speed $\boldsymbol{v}$ is determined, the momentum $\boldsymbol{p}$ cannot be determined." This is precisely the internal mechanism of the uncertainty principle, which not only ensures the continuity among things in reality, but also has been supported by a large number of experimental data when the principle was approved. In other words, the uncertainty principle and a large number of experimental data related to it have proven that the charge and static mass of electrons both are not constants.

7) In reality, along with the moving speed v of a high-speed particle (electron or photon) gradually approaching the c, the same particle exhibits two effects: on the one hand, according to the formula (5) its mass m keeps getting bigger ; on the other hand, its static mass m_0 keeps getting smaller due to the loss caused by electromagnetic radiation. The result of the final synthesis shows that its mass will follow along with its static mass gradually to approach zero. This is the energy shrinkage effect of high-speed particles, and also the primary factor causing the spectrum redshift. Therefore, the Big Bang theory is wrong. The nature of this matter is serious and of great significance, which is an objective reality that must be faced in modern physics, and should be verified to facilitate consensus.

8) As a light source of electromagnetic radiation, the speed of moving electrons in the storage ring, $v > 0.99c$ is sufficient to prove that the value c of light speed in vacuum is independent of the state of motion of the emitting body. On the contrary, it also proves that photons in reality must have static mass, and their size and moving speed are all different.

Conflict of Interest

There is no conflict of interest.

References

[1] Ding, J., 2022. The theory on thing's limits part 1: The norm of identifying truth. East African Scholars Journal of Psychology Behavioural Sciences. 4(4), 101-104. DOI: https://doi.org/10.36349/easjpbs.2022.v04i04.001

[2] Wen, H.W., 2009. Shi Lun Niu Dun Si Xiang Ti Xi Zhong De Zong Jiao Guan (Chinese) [Study on the Newtonian religious ideas]. Zhejiang Normal University: Jinhua. pp. 16. DOI: https://doi.org/10.7666/d.y1549479

[3] Einstein, A., 2000. Xiang Dui Lun (Chinese) [Relativity]. Hainan Press: Hainan. pp. 41-57.

[4] Ding, J., 2020. The research of using truth to restrict authoritative theories. Journal of Philosophy and Ethics. 2(1), 43-50.

[5] Newton, I., 2001. Zi Ran Zhe Xue De Shu Xue Yuan Li (Chinese) [The mathematical principles of natural philosophy]. Shaanxi Peoples Publishing House: Shaanxi. pp. 18.

[6] Sakyamuni. Chang A Han Jing (Chinese) [Dirghagama-sutra (Vol. 19)]. Sino-Culture Press: Beijing. pp. 609-610.

[7] de Broglie, L., 1992. Wu Li Xue Yu Wei Guan Wu Li Xue (Chinese) [Physics and microphysics]. The Commercial Press: Beijing. pp. 54.

[8] Einstein, A., 1977. Lun Dong Ti De Dian Li Xue (Chinese). [On the electrodynamics of moving bodies]. The collected works of Einstein (Vol. 2). The Commercial Press: Beijing. pp. 83-118.

[9] Ding, J., 2020. A brief analysis of the research

scheme of the cyclotron radiation from a single electron. Physics & Astronomy International Journal. 4(2), 60-64.
DOI: https://doi.org/10.15406/paij.2020.04.00202

[10] Ding, J., Hu, X.Q., 2018. Piercing the veil of modern physics: Part 1 & basics. Physics & Astronomy International Journal. 2(2), 128-134.
DOI: https://doi.org/10.15406/paij.2018.02.00074

[11] Jin, Y.M., 2001. Dian Zi Chu Cun Huan Wu Li (Chinese) [Electron storage ring physics]. USTC Press: Hefei. pp. 1.

[12] Wu, S.C., Wang, Z.Q., 1995. Jin Dai Wu Li Shi Yan (Chinese) [Modern physics experiment]. Peking University Press: Beijing. pp. 151.

[13] Ding, J., Hu, X.Q., 2014. The electro-ultimate particles and a new method for detecting the photon static mass. International Journal of Modern Physics and Application. 1(1), 1-8.

[14] Hawking, S.W., 1996. Shi Jian Jian Shi (Chinese) [A brief history of time]. Hunan Science & Technology Press: Changsha. p. 46-67.

BILINGUAL PUBLISHING GROUP

Journal of Electronic & Information Systems
https://journals.bilpubgroup.com/index.php/jeis

SHORT COMMUNICATION

A Doppler Location Method Based on Virtual Path Difference

Yu Tao

China Academy of Management Science, Beijing, 101100, China

ABSTRACT

This paper presents a Doppler passive location method for moving targets with fixed single station using the Doppler frequency shift and time difference information. First, based on the relationship between frequency shift and path difference, the virtual path difference is calculated from the measured value of Doppler frequency shift by means of mean value correction. Then, under the assumption that the target is moving at a constant speed along a straight line, two coaxial virtual double base arrays are constructed by using the moving track of the moving target based on the method of fixed period time difference. On this basis, the moving distance of the moving target can be calculated by using the ratio relationship between the frequency difference and the radial distance between the two adjacent detection points in the middle of the array, and the linear solution of the two double base path difference positioning equations. At this point, the relative coordinate position of the moving target can be obtained by directly using the linear solution of the double base path difference positioning equation again.

Keywords: Fixed single station; Passive location; Doppler frequency; Doppler frequency change rate; Frequency shift-path difference equation; Virtual path difference

1. Introduction

In electronic reconnaissance, when the signal bandwidth is insufficient, resulting in the ambiguity of time difference, or the phase calibration and time synchronization are limited by the system complexity, the location technology based on Doppler frequency shift information is preferred for target position estimation [1-7]. The above defects can be remedied by using this technology, and the requirements for receiving equipment are low. In addition, with the rapid development of frequency measurement technology and the continuous improvement of frequency measurement accuracy, the

*CORRESPONDING AUTHOR:
Yu Tao, China Academy of Management Science, Beijing, 101100, China; Email: 18001648180@163.com

ARTICLE INFO
Received: 24 February 2023 | Revised: 29 March 2023 | Accepted: 31 March 2023 | Published Online: 11 April 2023
DOI: https://doi.org/10.30564/jeis.v5i1.5496

CITATION
Tao, Y., 2023. A Doppler Location Method Based on Virtual Path Difference. Journal of Electronic & Information Systems. 5(1): 20-25. DOI: https://doi.org/10.30564/jeis.v5i1.5496

passive location technology based on the Doppler frequency shift has great development potential. However, due to the higher complexity of the nonlinear equation formed by Doppler frequency shift and target state compared with other positioning methods, this will not only lead to difficulties in analyzing the mathematical model, but also lead to the complexity of system design in practical engineering, so that the original advantages may be lost at last. At the same time, because the positioning system based on Doppler measurement is a nonlinear system, it is usually necessary to study the observability of positioning [8-10].

The author recently proposed a method that can directly convert the nonlinear Doppler frequency shift function into a linear solution in the research of the Doppler passive location of fixed targets by the motion detection station [11]. The new method obtains the relationship between the Doppler frequency shift and the radial path difference by differential processing of the radial velocity, so that the virtual path difference can be constructed according to the detected value of the Doppler frequency shift, and then the position of the target can be directly obtained by using the linear solution of the double-base path difference positioning equation. The new method not only has unique results, but also has very simple forms. The analysis complexity of passive location based on the Doppler frequency shift is effectively reduced.

Different from the application scenario of document [11], which is a way to detect fixed targets using motion detection stations, this paper studies the detection of moving target by fixed single station. Its difficulty is that the moving speed or moving distance of moving target is unknown. Obviously, its analysis is relatively more difficult. In this regard, the innovative method proposed by the author is: On the one hand, two coaxial virtual dual-basis arrays are constructed by using the moving track of the moving target, on the other hand, the moving distance of the moving target is calculated by using the linear solution of the two dual-basis path difference positioning equations according to the ratio relationship between the frequency difference and the radial distance between the adjacent detection points.

2. Path difference-frequency shift equation

2.1 Differential processing

According to the relationship between Doppler frequency shift and radial range change rate:

$$\frac{\partial r(t)}{\partial t} = v_r = v\cos\beta = \lambda f_d \tag{1}$$

where $r(t)$ is the radial distance; v_r is the radial velocity; v is the moving speed of the target; f_d is the Doppler frequency shift; λ is the wavelength; β is the leading angle.

For radial velocity, assuming that the change of time is short, the differential of distance to time can be converted into the ratio of path difference and time difference by using the difference calculation method:

$$\frac{\partial r(t)}{\partial t} \approx \frac{\Delta r}{\Delta t} \tag{2}$$

where Δr is the path difference; Δt is the time difference.

The geometric model corresponding to the mathematical model is shown in **Figure 1**. Substitute formula (2) for formula (1) to obtain the virtual path difference expression based on Doppler frequency shift measurement:

$$\Delta r = \lambda f_d \Delta t \tag{3}$$

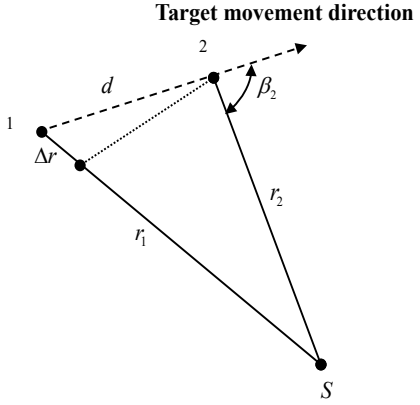


Figure 1. Geometric model of single motion station.

2.2 Mean value correction

If the Doppler frequency shift measurement value

of the moving target at position 1 is used, the expression of virtual path difference is:

$$\Delta r_{f1} = \lambda f_{d1} \Delta t \tag{4}$$

If the theoretical value of path difference is compared with the virtual path difference obtained based on frequency shift measurement:

$$\varepsilon = 100 \times \frac{\left| \Delta r - \Delta r_f \right|}{\Delta r} \% \tag{5}$$

where $\Delta r = r_1 - r_2$ is the theoretical value of path difference; Δr_f is the virtual path difference calculated according to equation (3).

The simulation calculation shows that the virtual path difference obtained has a large calculation error. However, further simulation calculation shows that if detection is also carried out at position 2 according to the geometric model in **Figure 1**, the Doppler frequency shift is obtained and used to calculate the path difference from position 1 to position 2, the virtual path difference obtained is:

$$\Delta r_{f2} = \lambda f_{d2} \Delta t \tag{6}$$

The subsequent simulation calculation shows that its calculation error is exactly opposite to the calculation error obtained by using the Doppler frequency shift at position 1. The upper and lower curves in **Figure 2** describe the relative calculation error curves of the virtual path difference obtained based on the Doppler frequency shift measurements at two locations.

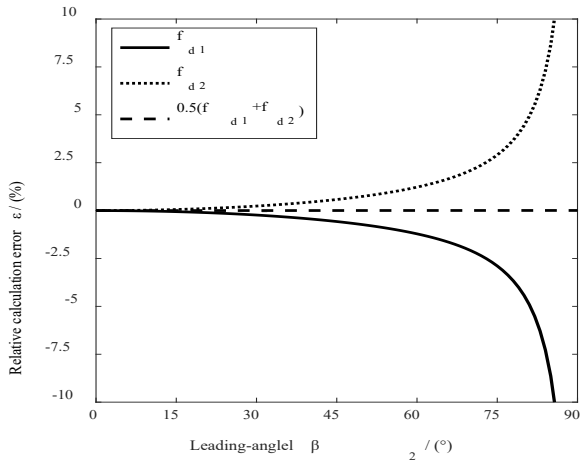


Figure 2. Mean value correction of path difference.

On this basis, if the average value of Doppler frequency shifts at two positions is used to calculate the virtual path difference:

$$\Delta r = 0.5\lambda \left(f_{d1} + f_{d2} \right) \Delta t \tag{7}$$

Then the calculation error is just offset. The middle curve in **Figure 2** shows that the relative calculation error of the virtual path difference obtained based on the Doppler average method tends to zero.

The parameters used in the simulation calculation are: $r_2 = 600\ km,\ d = 10\ km$.

3. Detection model

As shown in **Figure 3**, the target moves approximately at a uniform speed in a straight line, from position 1, through positions 2 and 3, to position 4. The fixed single station detects and receives the radiation signal from the target at a fixed time. If the fixed station detects four times continuously, then according to the linear solution of the double base array, two coaxial virtual double base arrays can be constructed simultaneously from the target's motion trajectory.

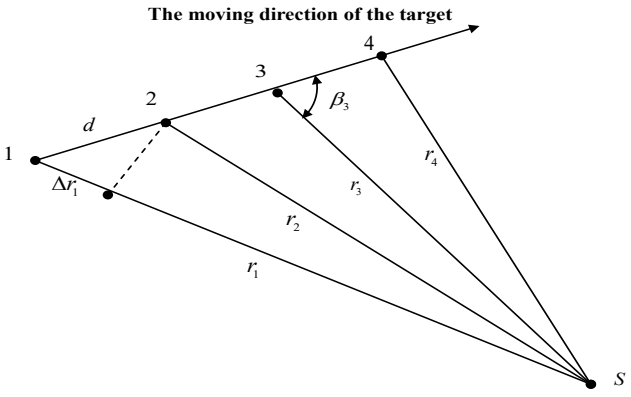


Figure 3. Geometric structure of virtual double-base array.

Based on the Doppler frequency shift measurement, three virtual path differences can be obtained:

$$\Delta r_1 = 0.5\lambda \left(f_{d1} + f_{d2} \right) \Delta t \tag{8}$$

$$\Delta r_2 = 0.5\lambda \left(f_{d2} + f_{d3} \right) \Delta t \tag{9}$$

$$\Delta r_3 = 0.5\lambda \left(f_{d3} + f_{d4} \right) \Delta t \tag{10}$$

Then, two ranging solutions are directly given by the double base path difference ranging solution [12]:

$$r_2 = \frac{2d^2 - \Delta r_1^2 - \Delta r_2^2}{2(\Delta r_1 - \Delta r_2)} \tag{11}$$

$$r_3 = \frac{2d^2 - \Delta r_2^2 - \Delta r_3^2}{2(\Delta r_2 - \Delta r_3)} \tag{12}$$

4. Moving distance of target

4.1 Ratio of Doppler change rate of adjacent nodes

According to the definition of Doppler frequency shift change rate and the target motion trajectory, the following two equations can be listed at two adjacent detection points in the middle of the virtual array:

$$\dot{f}_{d2} = \frac{v_{t2}^2}{\lambda r_2} \tag{13}$$

$$\dot{f}_{d3} = \frac{v_{t3}^2}{\lambda r_3} \tag{14}$$

where $\dot{f}_{d2}$ and $\dot{f}_{d3}$ are Doppler change rates; λ is the wavelength. r_2 and r_3 are radial distances; v_{t2} and v_{t3} are tangential velocities.

The ratio of the Doppler frequency shift change rate between these two adjacent detection points is:

$$\frac{\dot{f}_{d3}}{\dot{f}_{d2}} = \frac{r_2}{r_3}\frac{v_{t3}^2}{v_{t2}^2} \tag{15}$$

According to the sine theorem, the ratio between the radial distances of the two adjacent detection points is:

$$\frac{r_2}{r_3} = \frac{\sin\beta_3}{\sin\beta_2} = \frac{v\sin\beta_3}{v\sin\beta_2} = \frac{v_{t3}}{v_{t2}} \tag{16}$$

where β_1 and β_2 are leading angles.

Replace it with formula (15) to get:

$$\frac{\dot{f}_{d3}}{\dot{f}_{d2}} = \frac{r_2^3}{r_3^3} \tag{17}$$

4.2 Frequency difference ratio

Use differential processing:

$$\dot{f}_d = \frac{\Delta f_d}{\Delta t} \tag{18}$$

where Δf_d is the Doppler frequency difference.

Moreover, due to the constant range or periodic detection, the adjacent time difference is nearly equal, so the ratio of Doppler change rate between adjacent nodes can be approximately expressed by the ratio of Doppler frequency difference:

$$q = \frac{\dot{f}_{d3}}{\dot{f}_{d2}} \approx \frac{\Delta f_{d32}}{\Delta f_{d21}} = \frac{f_{d3} - f_{d2}}{f_{d2} - f_{d1}} \tag{19}$$

where f_{d1}, f_{d2} and f_{d3} are Doppler frequency shifts.

4.3 Solution of the equation

Substitute the two ranging solutions (11) and (12) into the frequency difference ratio:

$$\sqrt[3]{q} = \frac{\left(2d^2 - \Delta r_1^2 - \Delta r_2^2\right)}{\left(\Delta r_1 - \Delta r_2\right)}\frac{\left(\Delta r_2 - \Delta r_3\right)}{\left(2d^2 - \Delta r_2^2 - \Delta r_3^2\right)} \tag{20}$$

After sorting, there are:

$$\sqrt[3]{q}\frac{\left(\Delta r_1 - \Delta r_2\right)}{\left(\Delta r_2 - \Delta r_3\right)}\left(2d^2 - \Delta r_2^2 - \Delta r_3^2\right) = \left(2d^2 - \Delta r_1^2 - \Delta r_2^2\right) \tag{21}$$

Finally, it can be solved as follows:

$$d = \sqrt{\frac{\left(\Delta r_1^2 + \Delta r_2^2\right) - A\left(\Delta r_2^2 + \Delta r_3^2\right)}{2(1-A)}} \tag{22}$$

including:

$$A = \sqrt[3]{q}\frac{\left(\Delta r_1 - \Delta r_2\right)}{\left(\Delta r_2 - \Delta r_3\right)}$$

5. Location of the target

5.1 Distance

After the moving distance of the moving target is obtained, the distance between the detection station and the moving target can be obtained from the double base path difference solution again:

$$r_2 = \frac{2d^2 - \Delta r_1^2 - \Delta r_2^2}{2(\Delta r_1 - \Delta r_2)} \tag{23}$$

Figure 4 shows the relative calculation error of the ranging solution at a different radial distances and different motion distances. It can be seen that the longer the radial distance, the shorter the moving

distance of the target, and the smaller the relative calculation error. Divergence will occur when it approaches 90 degrees.

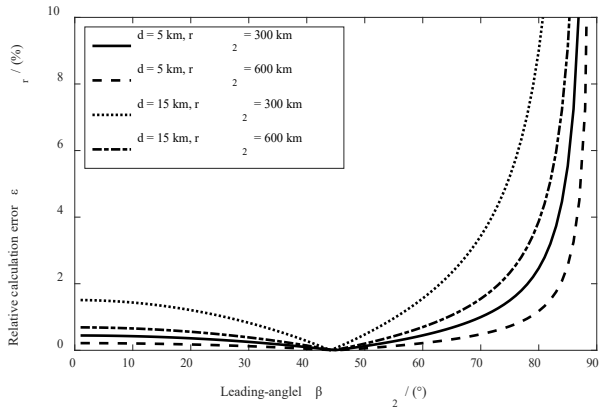


Figure 4. Relative calculation error of ranging solution.

5.2 Orientation of the target

Using two virtual path differences, the included angle between the moving direction of the target and the radial distance can be directly obtained from the double base path difference DF solution [12]:

$$\cos\beta_2 = \frac{\left(d^2 - \Delta r_1^2\right)\Delta r_2 + \left(d^2 - \Delta r_2^2\right)\Delta r_1}{d\left(2d^2 - \Delta r_1^2 - \Delta r_2^2\right)} \tag{24}$$

Figure 5 shows the relative calculation error of the DF solution at different radial distances and different motion distances.

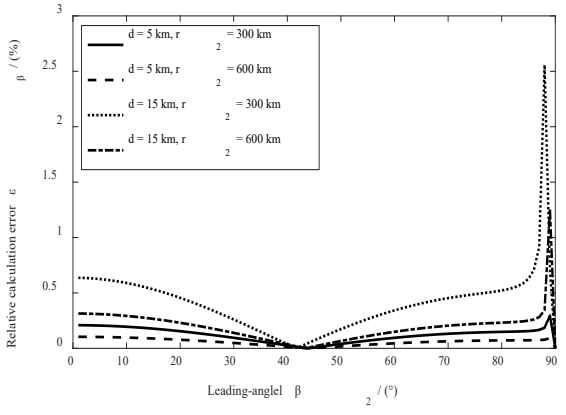


Figure 5. Relative calculation error of DF solution.

5.3 Target's movement speed

After the azimuth angle is obtained, the moving speed of the target can be directly solved by Doppler frequency shift:

$$v = \frac{\lambda f_{d2}}{\cos\beta_2} \tag{25}$$

where v is the moving speed of the target.

Or the moving speed of the target in the detection time can be approximately obtained from the moving distance and detection time of the target:

$$v = \frac{1}{\Delta t}\sqrt{\frac{\left(\Delta r_1^2 + \Delta r_2^2\right) - A\left(\Delta r_2^2 + \Delta r_3^2\right)}{2\left(1 - A\right)}} \tag{26}$$

6. Conclusions

Based on the innovative results of the basic application theory of passive location technology [3], the Doppler location method using virtual path difference presented in this paper greatly simplifies the system design. In fact, the author's research has shown that the path difference-frequency shift equation will help to construct new passive location methods.

The existing mathematical description of the Doppler frequency shift is basically carried out in one-dimensional space. The author's earlier research proved that when the azimuth between the wave source and the observer changes with time, the Doppler shift should be a function on the two-dimensional plane [13]. The Doppler frequency shift on the two-dimensional plane can always be decomposed into the sum of two terms, one of which is only related to the radial velocity, and the other is related to the radial acceleration. In addition, the given formula includes two Doppler frequency shifts at different times, which provides a mathematical method for obtaining the Doppler frequency shift value at the current time by using the Doppler frequency recurrence at the previous time.

The recent research results [11] show the relationship between path difference and frequency shift, which provides a new solution for determining the position of the target directly using Doppler frequency shift. The frequency shift-path difference equation obtained by mean value correction is the analysis

basis of the new method proposed in this paper. It is based on the modified frequency shift-path difference equation with a small relative calculation error that can directly use the double base path difference positioning theory to obtain the results with good calculation accuracy.

With regard to the relationship between frequency shift and path difference, the author's current concern is: Can we derive a more rigorous mathematical formula? Can it rise to the height of a physical equation?

Conflict of Interest

There is no conflict of interest.

References

[1] Zhou, Ch., Ma, C.Sh., Ying, T., et al., 2022. Source localization using Doppler frequency shift with erroneous carrier frequency. Journal of University of Electronic Science and Technology of China. 51(4), 529-534.
DOI: https://doi.org/10.12178/1001-0548.2022050

[2] Tian, M.H., Ma, M., Zhang, W.Y., 2018. Target localization based on Doppler frequency and range sum. Journal of Terahertz Science and Electronic Information Technology. 16(6), 984-988.
DOI: https://doi.org/10.11805/TKYDA201806.0984

[3] Wang, D., Yin, J.X., et al., 2017. Direct iocalization method for constant modulus source based on Doppler frequency shifts. Acta Aeronautica ET Astronautica Sinica. 38(9), 284-297.
DOI: https://doi.org/10.7527/S1000-6893.2017.321084

[4] Tao, Y., 2011. Airborne Doppler direct ranging method based on angle change rate. Advances in Aeronautical Science and Engineering. 2(3), 335-338.
DOI: https://doi.org/10.3969/j.issn.1674-8190.2011.03.018

[5] Tao, Y., 2012. A moving single station Doppler passive location method. Guidance & Fuze. 33(1), 16-18,32.
DOI: https://doi.org/10.3969/j.issn.1671-0576.2012.01.004

[6] Zhou, Zh., Wang, G.Ch., 2008. Passive location and tracking of maneuvering targets by airborne single station. Electronics Optics & Control. 15(3), 60-63.
DOI: https://doi.org/10.3969/j.issn.1671-637X.2008.03.016

[7] Liu, C.F., 2011. Wu Yuan Ding Wei Yu Gen Zong (Chinese) [Passive Location and Tracking]. Xidian University Press: Xi'an.

[8] Xiang, F.H., Wang, J.G., 2021. Observability and simulation analysis of fixed single observer passive location. Modern Defense Technology. 49(04), 72-78.

[9] Zhan, R.H., Wang, L., Wan, J.W., 2020. Observable conditions for passive location and tracking of maneuvering targets based on azimuth and Doppler. Journal of National University of Defense Technology. 29(1), 54-58.

[10] Deng, X.P., 2007. Comments on observability of single observer passive location. Strategic Study of CAE. 11, 54-62.

[11] Tao, Y., 2020. An airborne passive positioning method based on angle and frequency difference measurement. Proceedings of the 2020 4th International Conference on Electronic Information Technology and Computer Engineering (EITCE 2020); 2020 Nov. p. 296-301.
DOI: https://doi.org/10.1145/3443467.3443771

[12] Tao, Y., 2017. Wu Yuan Tan Ce Ding Wei Ji Shu (Chinese) [Technology of passive detection location]. National Defense Industry Press: Beijing.

[13] Tao, Y., 2010. Expression for Doppler shift in two-dimensional plane. The 2010 International Conference on Information, Electronic and Computer Science. Scientific Research Publishing: USA. p. 1-4.

BILINGUAL PUBLISHING GROUP

Journal of Electronic & Information Systems
https://journals.bilpubgroup.com/index.php/jeis

ARTICLE

An Improved Power Efficient Clock Pulsed D Flip-flop Using Transmission Gate

B.Syamala, M.Thamarai*

Electronics and Communication Engineering, Sri Vasavi Engineering College, AP, 534101, India

ABSTRACT

Recent digital applications will require highly efficient and high-speed gadgets and it is related to the minimum delay and power consumption. The proposed work deals with a low-power clock pulsed data flip-flop (D flip-flop) using a transmission gate. To accomplish a power-efficient pulsed D flip-flop, clock gating is proposed. The gated clock reduces the unnecessary switching of the transistors in the circuit and thus reduces the dynamic power consumption. The clock gating approach is employed by using an AND gate to disrupt the clock input to the circuit as per the control signal called Enable. Due to this process, the clock gets turned off to reduce power consumption when there is no change in the output. The proposed transmission gate-based pulsed D flip-flop's performance with clock gating and without clock gating circuit is analyzed. The proposed pulsed D flip-flop power consumption is 1.586 μw less than the without clock gated flip-flop. Also, the authors have designed a 3-bit serial-in and parallel-out shift register using the proposed D flip-flop and analyzed the performance. Tanner Electronic Design Automation tool is used to simulate all the circuits with 45 nm technology.

Keywords: Pulsed D flip-flop; Clock gating; Low power; Shift register; Transmission gate

1. Introduction

The main concern with VLSI circuits is achieving low power and compact battery size with a long life. Area, power, speed, and performance are all critical parameters for VLSI designers to consider while optimizing their designs [1-4]. Flip flops are the major building blocks of digital circuits and also consume around 80% of the total power consumption in the circuits. Therefore, reducing the power consumption of flip-flops can signifi-

*CORRESPONDING AUTHOR:
M.Thamarai, Electronics and Communication Engineering, Sri Vasavi Engineering College, AP, 534101, India; Email: thamarai.muthusamy@srivasaviengg.ac.in

ARTICLE INFO
Received: 26 March 2023 | Revised: 7 April 2023 | Accepted: 10 April 2023 | Published Online: 14 April 2023
DOI: https://doi.org/10.30564/jeis.v5i1.5574

CITATION
Thamarai, M., Syamala, B., 2023. An Improved Power Efficient Clock Pulsed D Flip-flop Using Transmission Gate. Journal of Electronic & Information Systems. 5(1): 26-35 DOI: https://doi.org/10.30564/jeis.v5i1.5574

cantly reduce the power consumption of digital systems [5]. Any circuit's overall power consumption will include both static and dynamic power. In VLSI circuits, power gating, clock gating, adiabatic method, and other approaches are utilized to reduce static and dynamic power consumption. One of the most common strategies for power optimization in CMOS devices is clock gating [1]. Almost 50% of the dynamic power is consumed by the clock distribution network in processors and clock gating can be used at several levels of the hierarchy in VLSI design and almost 50% of the dynamic power is consumed by the clock distribution network in integrated circuits. It is feasible to turn off the clock signal that drives a large functional unit, lowering power consumption on both its internal nodes and the clock line.

Pulsed Flip-Flops (P-FF): Pulsed flip-flops are used in high speed circuits and take an important role in power reduction. Short duration pulses are generated and used to trigger the flip-flops instead of level or edge triggering. P-FFs consist of a single latch structure and a clock pulse generator (PG). If the width of the clock pulse is sufficiently narrow, the P-FF acts like a Master-Slave Flip-Flops (MS-FF) with less timing overhead. Jhon K et al. showed that several pulse triggered flip-flop architecture for high speed allocations [5]. PF-FFs can be classified as implicit and explicit type based on the way of pulse generation. In implicit style, the pulse generation part is inbuilt into the latch structure and in explicit it is external to the latch structure. In explicit pulsed flip-flops, the pulse generation structure can be common among neighboring flip-flops and saves chip area.

Transmission Gate: The parallel connection of nmos and pmos forms the Transmission Gate (TG). The transmission gate's on or off condition is controlled by the nmos and pmos gate inputs. The gate is in on condition when nmos transistor gate gets Clk input and pmos transistor gate gets Clkb input and TG the circuit is shown in **Figure 1**.

TGs are bilateral switches since the input and output are interchangeable. TGs are used to form D latch and D flip-flops in logic circuits. The positive D latch using the transmission gate is shown in **Figure 2**.

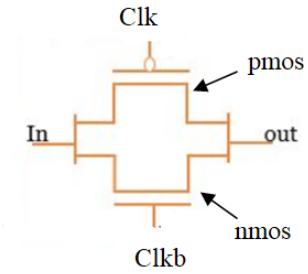


Figure 1. Transmission gate.

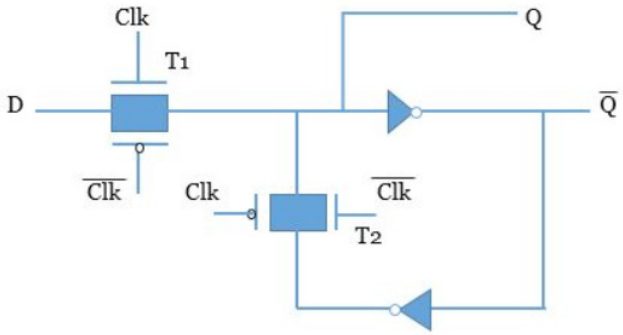


Figure 2. Transmission gate based D latch.

It consists of two TGs and two inverters. When Clk = high, T1 is ON and T2 is OFF, so output (Q) directly follows the input (D). Otherwise, the output is the stored values in the latch.

The proposed work deals with transmission gate based pulsed D flip-flop design with clock gating for power reduction. The paper is organized as follows. Section 2 describes the related work and Section 3 describes the proposed clock gated transmission gate based D flip-flop and Section 4 describes the implementation of a shift register using the proposed D flip-flop. Section 5 discusses the results and Section 6 gives the conclusions.

2. Related work

Ultra-low power pulse-triggered carbon nanotube field-effect transistor (CNTFET) based flip-flop is proposed by Karim et al. [6] for low power applications. The author used signal feed through technique and also optimized the discharging path to reduce the delay time for the '1' to '0' transition. The proposed

design is also area efficient, since the number of transistors used is less when compared to the other pulse triggered flip-flop schemes. An explicit type double edge pulse-triggered flip-flop is designed by Singh et al. [7]. The author used Exor gate based pulse generator to obtain a short pulse at the rising edge and falling edge of the clock signal. The designed flip-flop has less power delay product and is more suitable for high speed applications. Phyu et al. designed the pulse triggered flip-flop using static latch and dynamic pulse generator [8]. The author used explicit type double edge triggering approach. Ameena et al. proposed another explicit type pulse triggered flip-flop based on data feed through scheme [9]. The proposed data feed through scheme reduces the long discharge path in conventional explicit pulse triggered flip-flops and is also suitable for high speed applications. John K. [10] proposed an implicit type pulse triggered flip-flop by making conditional enhancement in width and height of the triggering pulses by using an additional pmos transistor in the structure for high speed applications. J.F.Lin et al. [11] proposed an explicit type pulsed flip-flop by using True single-phase clock latch based on a signal feed-through scheme.

Karimni [12] proposed an ultra low pulse triggered flip-flop with optimized leakage power. The author used transmission gate to control the input data and the leakage power. Also, the Pulse Generator (PG) is modified to reduce the number of required transistors and the clock pulse delay. Panahifar, E. & Hassanzadeh designed a signal feed-through flip-flop that uses a pass transistor to feed input data directly to the output and minimizes the dynamic power [13]. Another efficient high speed conditional feed through pulsed flip-flop was proposed by Pan.D et al. [14]. Pulsed flip-flop using a transmission gate is proposed by Prakalya et al. [15]. They used current mode distribution of clock pulses than the conventional voltage mode operation. Consoli, E. et al. designed a modified transmission gate based Master Slave Flip-Flop (MS F/F) to minimize the delay in the classical transmission gate based approach [16]. Islam R. et al. [17] proposed a current mode pulsed flip-flop combined with a current mode signaling transmitter, which reduces the power consumption of clock distribution networks more than the voltage mode distribution of clock signal.

Clock gating is one of the efficient methods in clock power reduction while the input signal switching activity is low [18,19]. John K et al. discussed the effect of clock gating on conditional pulsed flip-flops and analyzed in their work [20]. Sharma, D. K designed and analyzed different clock gating circuits such as latch based, flip-flop based and gate based [21]. The author analyzed the circuits in terms of delay, power and area and proved that gate based clock gating circuits are more power efficient [21]. A latch-based clock gating technique is used in the design of SRAM and sequential counter and verified the power reduction in the work [22].

3. Proposed work

3.1 Clock gating

The static (leakage) power grows dramatically with each generation of technology, dynamic power still dominates the total power dissipation of general-purpose microprocessors [2]. The reduction in transistor size, interconnect optimization, applying gated clock, applying variable threshold voltages, and the management of dynamic supply voltage are all effective circuit strategies for reducing dynamic power usage. Clock gating is one of the efficient methods in clock power reduction when the input signal switching activity is low.

The clock gating approach needs an additional logic circuit to generate a clock enabling signal [19,21,23]. The logic circuit compares the input and output of the latch and if there is any change in the output, which will be enabled only when the need for output changes to logic 0 or 1 value according to the circuit design. The clock gating approach is used to prevent unnecessary clock pulses from being sent to the circuit when there is no need for output change. For example, already flip-flop has output 1, if the next input bit is 1 ten no need to apply the clock pulse to the flip-flop. This will reduce the unnecessary

switching of the circuit capacitance and thus reduces the dynamic power consumption. Moreover, the toggling nature of the clock consumes more power in registers.

In flip-flops, the switching activity is increased due to the transition from 0 to 1 and 1 to 0, and they require more power. The first step in implementing clock gating is to identify the location (module) which consumes large power. The next step is to study the switching activity of the module and based on that designing of a control logic circuit that produces the enable signal. The enable signal is used to turn off and on the clock signal which is going to be applied at that location. When there is no change in the output of a particular register or flip-flop, then that unit does not require a clock pulse, since the unit is going to maintain the same output. The clock gating approach cuts the clock pulse to that particular register or flip-flop and thus in turn saves the dynamic power dissipation. In a sequential circuit, the simple solution is to cut the clock pulses which are not going to change the output state of the circuit. This reduces power consumption. The clock gating to the flip-flop block diagram is shown in **Figure 3**.

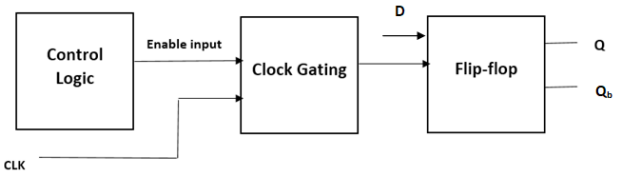


Figure 3. Clock gating technique to the flip-flop.

3.2 Flow chart for clock gating

The flowchart shown in **Figure 4** shows the simple clock gating method. The clock pulses are applied to the circuit whenever the clock and enable input both are in logic '1' condition. If the enable signal is logic '0' then the clock pulse is blocked to the circuit. Clock gating is a simple structure and easy to implement. Any of the following gates can be utilized in a gated clock gating approach. AND, OR, and NOR gates are the three gates that are used in clock gating. Only when both inputs reach a high value does the AND gate on. Because of the short transition period from high to low, when enabling pin = 1, the counter in negative edge triggering type increments by one count. When the enable pin = 1 during the positive edge of the clock, there is a long fall time, which results tiny glitches (short duration clock pulses) and the circuit produces output with an error. The AND gate is mostly used for activating negative edge triggering circuits.

There are two inputs, Enable and Clock are given to the AND gate. The circuit operation is as follows. When the Enable is low and the Clock is low then the transistor T1, T4, and T6 are in ON condition and the T2, T3, T5 are in OFF, so the output is low. When the Enable is low and the Clock is high the transistors T1, T3, and T6 are in ON condition the transistor T2, T4, and T5 are in OFF condition, so the output is low. Similarly, when the Enable is high and the Clock is low, the transistors T2, T4, and T6 are in ON condition and the transistor T1, T3, and T5 are in OFF condition and the output is low. When the Enable is high and the Clock is high the transistor T2, T3, and T5 are in ON condition and the transistors T1, T4, and T6 are in OFF condition and the output is high. The clock gating circuit output is shown in **Figure 6**. Whenever the Enable and the Clock both are set to HIGH then the output is HIGH and the gated clock is given to the proposed D flip-flop.

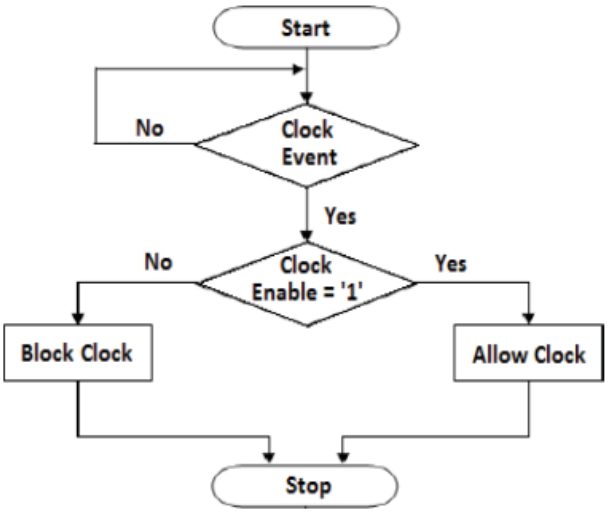


Figure 4. Flowchart of clock gating.

The proposed clock gating circuit is shown in **Figure 5**.

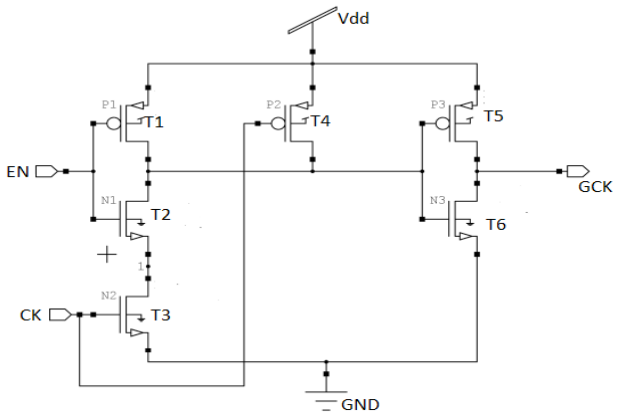


Figure 5. Schematic of AND clock gating.

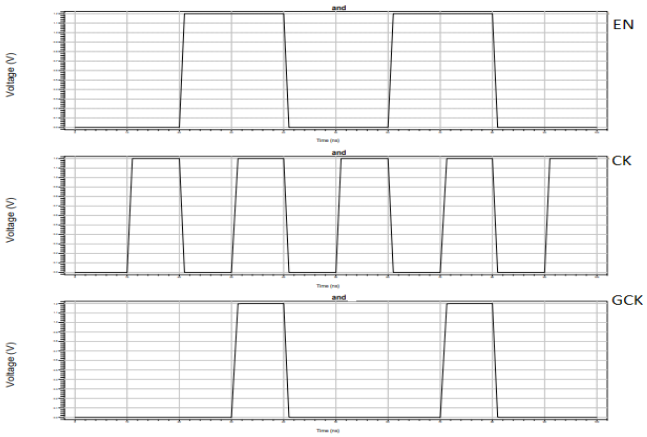


Figure 6. Waveforms of AND Clock gating.

3.3 Transmission gate based D flip-flop

Transmission gate logic is one of the additional ways to reduce power consumption by using more power-efficient circuitry to implement the D flip-flop [8,24]. The transmission gate logic is used to construct a flip-flop register stage and is also used at the input stage in memristors [25]. The **Figure 7** shows a leading-edge triggered D flip-flop without Set, Reset, or Clear signals with the clock signal. D is the input signal and Clk and Clkb denote clock and inverted clock signals respectively.

The D pin is used to provide input in a flip-flop. Transmission gates are equipped with clock pulses that make the transmission gates on and off. When the Clk = 1 and Clkb = 0, then the conditions of TGs are as follows. TG1 = Off, TG2 = On, TG3 = On and TG4 = Off and the data stored in the latch (N1) will be transferred to the output. When Clk = 1 and Clkb = 0, then TG1 = On, TG2 = Off, TG3 = Off and TG4 = On and the data at D will be transferred to the first latch (node N1). The circuit acts as a master-slave flip-flop. The modified flip-flop circuit using an efficient clock distribution network is shown in **Figure 8** [15].

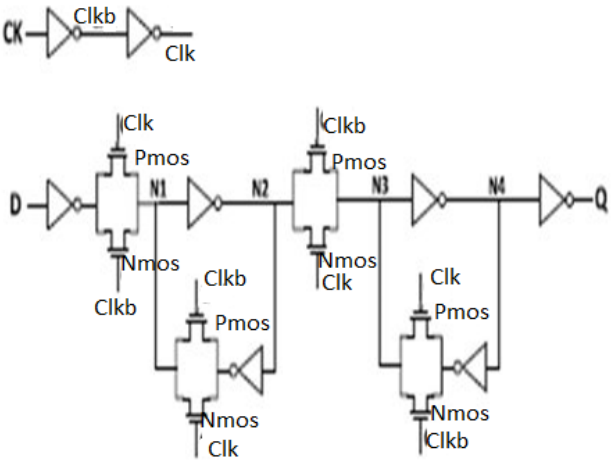


Figure 7. Schematic of transmission gate-based D flip-flop.

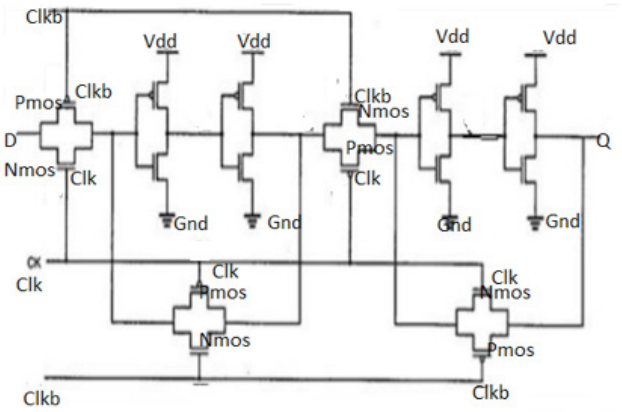


Figure 8. Schematic of TG-based D flip-flop with the clock distribution network.

The proposed flip-flop with the gated clock is shown in **Figure 9**. Whenever the enable and the clock are set to HIGH then only the output is HIGH. The gated clock is given to the proposed D flip-flop. As in **Figure 9**, the gated Clk is applied to the inverter to obtain the Clk and Clkb signals.

In the proposed circuit D flip-flop is implemented using 24 transistors. The clock gating logic is added to the D flip-flop, which increases the complexity of the circuit, but the power consumption is reduced be-

cause of the AND clock gating logic. The transistors T1 and T2 combined to form a transmission gate i.e., TG1. The transistors T5 and T6 combined to form a transmission gate i.e., TG2 and the transistors T9 and T10 combined to form a transmission gate named TG3. The transistors T13 and T14 form transmission gate, named TG4. If D input is High and the clock is High, then the gated clock output will be one. That is Clk = 1 and Clkb = 0. Now, the TG1 and TG4 will be in on condition and TG2 and TG3 will be in off condition. The input D is transmitted to the first latch and the output is the data stored in the second latch. When the gated clock is zero, that is Clk = 0 and Clkb = 1, then the TG1 and TG4 will be in off condition and TG2 and TG3 will be in on condition and the data in the first latch is moved to the second latch, and whenever clock one, the data will be made available at the output.

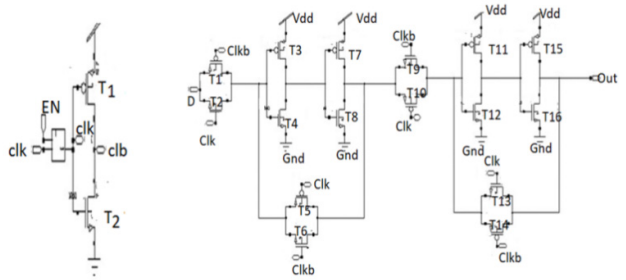


Figure 9. Schematic of proposed transmission gate based D flip-flop with clock gating.

Figure 10 shows the proposed D flip-flop output waveforms with clock gating.

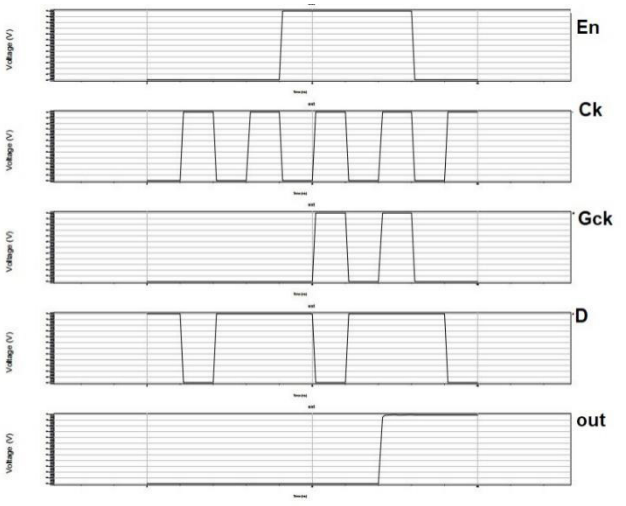


Figure 10. Waveforms of TG-based D flip-flop with AND clock gating.

4. Serial in parallel out shift register using proposed D flip-flop

A 3-bit serial-in and parallel-out shift register is designed using the proposed pulsed transmission gate based D flip-flop with clock gating. Shift registers consist of three proposed D flip-flops. Each flip-flop module is with clock gating circuit. Clock pulses and enable signal are applied to all the flip-flops and the data stored in the flip-flops are moved based on the gated clock output. The gated clock becomes high (logic '1'), when both the Clk and enable input are high. The shift register circuit is shown in **Figure 11**. The inputs D and EN (Enable) and CK (Clk) are given as follows.

input D:	1 1 1 0 1 1 1 0 1 1
input EN:	1 1 0 0 1 1 0 0 1 1.
Clock:	0 1 0 1 0 1 0 1 0 1
Gated clock output:	0 1 0 0 0 1 0 0 0 1

The data shifting with the gated clock condition (En=1 & Clk=1) is shown in **Figure 12**.

As shown **Figure 12**, V(3) becomes '1' during the first gated clock pulse and V(2), V(1) becomes '1' during the second and third gated clock pulses.

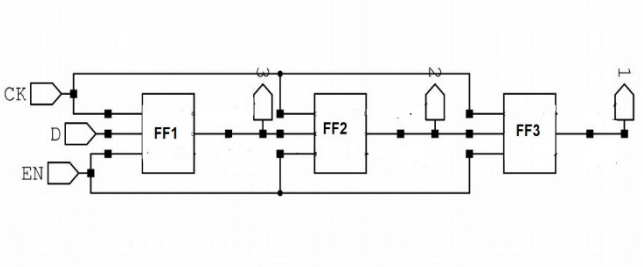


Figure 11. Schematic of three bit shift register with clock gating.

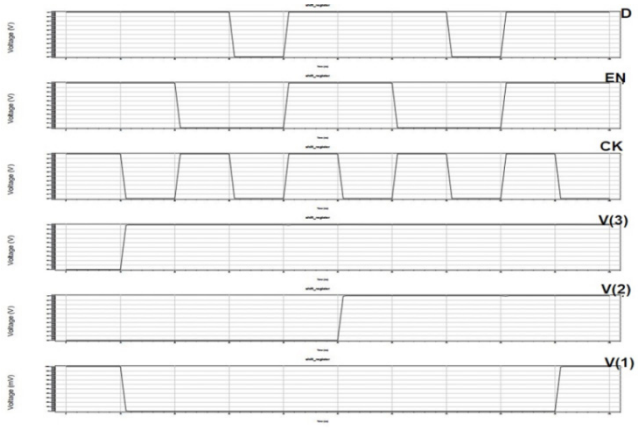


Figure 12. Waveforms of three-bit shift register.

5. Implementation results in tanner tool

The proposed transmission gate based D flip-flop with clock gating circuit and shift register circuits discussed in Sections 3 and 4 are simulated and their performance is measured using tanner tool with 45 nm technology.

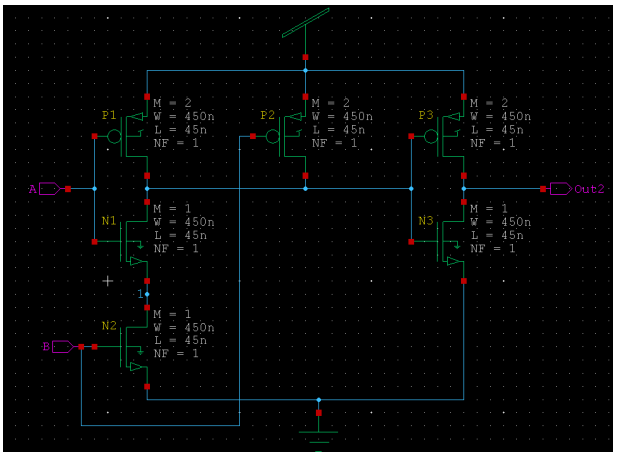


Figure 13. Schematic of AND clock gating in tool.

Figure 13 shows the AND gate clock gating circuit implementation of the schematic shown in **Figure 5**. In the circuit, the input terminal A refers to the enable input and B refers the clock input. The proposed transmission gate based D flip-flop with clock gating schematic in tool is shown in **Figure 14**, which is the same as the circuit in **Figure 9**.

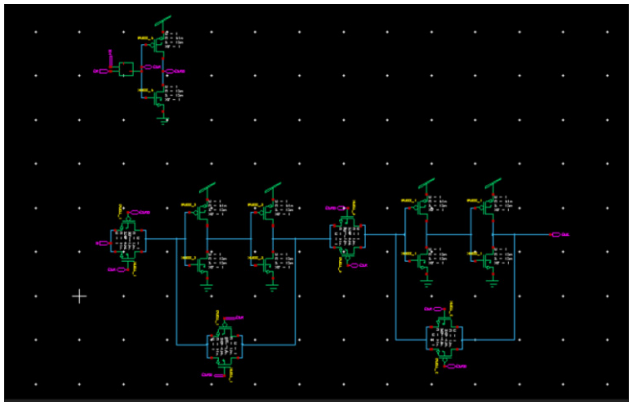


Figure 14. Schematic of proposed TG-based D flip-flop with clock gating in tool.

Figure 15 shows the output of the proposed transmission gate based D flip-flop with clock gating. Whenever the Enable and the Clock is in high condition then only the output data are transferred otherwise the output remains in the same condition. This reduces power consumption. The first waveform shows the Enable input and the second one is the clock input. The gated clock input is shown in the third waveform. The D input and flip-flop outputs are shown in waveforms 4 and 5 respectively.

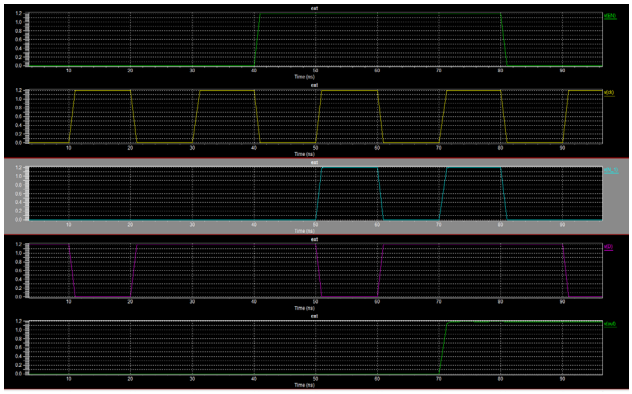


Figure 15. Output waveforms TG-based D flip-flop with clock gating in tool.

The inputs and obtained output values are given below.

Enable: 0 0 0 0 1 1 1 1 0 0 (input)
Clk: 0 1 0 1 0 1 0 1 0 1 (input)
gated clock: 0 0 0 0 0 1 0 1 0 0 (output based on clock gating)
D input: 1 0 1 1 1 0 1 1 1 0 (input)
Vout: 0 0 0 0 0 0 0 1 1 1 (flip-flop output)

The flip-flop output is zero for the first five D inputs, since the gated clock is zero. The output remains in its previous condition (0 value). Next gated clock becomes 1 and the D value is 0 and is stored in the flip-flop (output becomes 0). It is maintained until the gated clock becomes 1. The gated clock becomes 1, the D input is also 1 (8th bit), the flip- flop output changes to 1.

Figure 16 shows the serial in parallel out shift register using the proposed D flip-flop with clock gating. The register accepts serial input (one bit at a time through a single data line) and generates parallel output. The circuit consists of three connected D flip-flops [21]. All three of the flip flops are coupled with the clock signal and the enable signal and D input signal. The first flip-flop output is given to the

second flip-flop as D input and so on. All the flip-flop receives the same clock signal and is having a clock gated module. The gated logic circuit allows gated clock only whenever there is a change in input value.

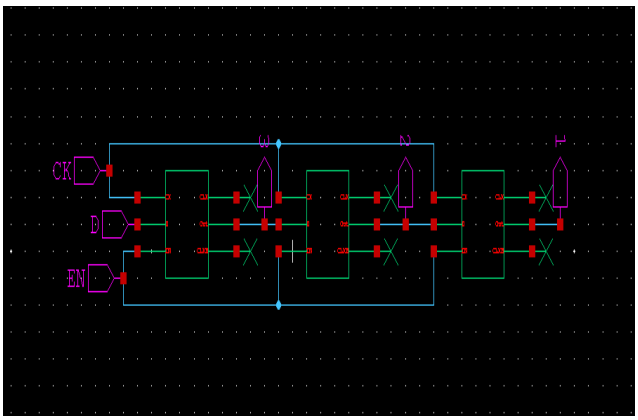


Figure 16. Serial-in parallel-out shift register in tool.

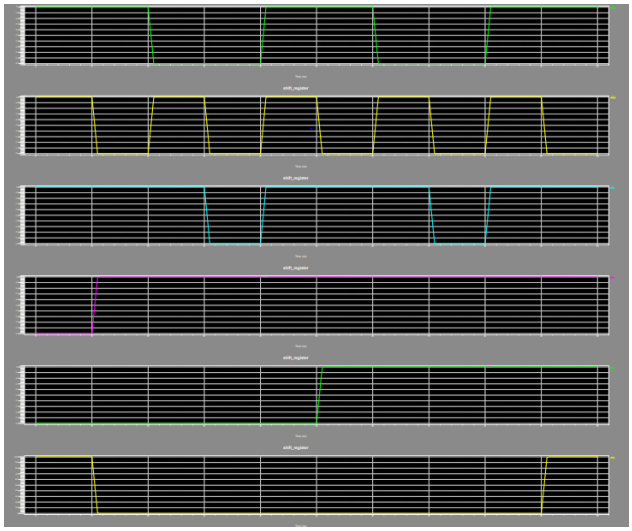


Figure 17. Output waveforms of the shift register in tool.

Figure 17 shows the output of the three-bit shift register and five clock pulses are applied to the circuit to store the data '1 1 1'. The gated pulse triggering circuit (clock gating) allows only three pulses to the flip-flop. These pulses are used to move and store the given input D values in the shift register. This avoids unnecessary switching of transistors in the circuit and saves power.

The power consumption of the proposed D flip-flop with clock gating is less when compared to the existing transmission gate based flip-flops and is given in **Table 1**.

Table 1. Comparison of proposed work and existing work.

Method	Power	Number of transistors	Delay
Transmission gate-based D flip-flop without clock gating (Conventional)	0.24 mW	24	20.02 ns
Transmission gate based D flip-flop with clock gating	0.17 mW	30	10.23 ns
Modified Transmission gate-based D flip-flop without clock gating	4.699 µW	18	60.02 ns
Modified Transmission gate-based D flip-flop with clock gating (Proposed work)	3.092 µW	24	49.16 ns

6. Conclusions

In this paper, a clock gated pulsed D flip-flop using a transmission gate was implemented. The gated clock reduces the unnecessary switching of the transistors in synchronous circuits and in turn, the dynamic power consumption of the circuit is also reduced. First, the clock gating circuit was implemented. Next, the transmission gate-based D flip-flop, modified transmission gate-based D flip-flop performances are measured with and without clock gating. The performances are analyzed in terms of power, the number of transistors and delay. The power reduction of 1.586 µw (34%) was achieved using the clock-gated modified transmission gate-based D flip-flop when compared to the without clock gating condition of the same circuit. The delay is more in the proposed circuit, since the gated clock circuit increases the delay. A 3-bit serial-in and parallel-out shift register was also constructed using the proposed D flip-flop and its average power was 4.22 µw watts which is less than the conventional 3-bit shift register without clock gating. The proposed design pulsed D flip-flop is useful for low power applications.

Conflict of Interest

There is no conflict of interest.

References

[1] Attaoui, Y., Chentouf, M., Ismaili, Z.E.A.A., et al., 2021. Clock gating efficiency and impact on power optimization during synthesis flow. 2021 International Conference on Microelectronics (ICM); 2022 Jan 07. New Cairo City. New York: IEEE. p. 13-16.

[2] Attia, K.M., El-Hosseini, M.A., Ali, H.A., 2017. Dynamic power management techniques in multi-core architectures: A survey study. Ain Shams Engineering Journal. 8(3), 445-456.

[3] Natarajan, V., Nagarajan, A.K., Pandian, N., et al., 2018. Low power design methodology. Very-large-scale integration. IntechOpen: London. pp. 47.

[4] Varadharajan, S.K., Nallasamy, V., 2017. Low power VLSI circuits design strategies and methodologies: A literature review. 2017 Conference on Emerging Devices and Smart Systems (ICEDSS); 2017 Mar 03-04. Mallasamudram. New York: IEEE. p. 245-251.

[5] John, K., RS, V.K., Kumar, S.S., 2018. Performance comparison of explicit pulsed flip-flops in low power and high-speed applications. 2018 International Conference on Circuits and Systems in Digital Enterprise Technology (ICCSDET); 2018 Dec 21-22. Kottayam. New York: IEEE. p. 1-6.

[6] Karimi, A., Rezai, A., Hajhashemkhani, M.M., 2019. Ultra-low power pulse-triggered CNTFET-based flip-flop. IEEE Transactions on Nanotechnology. 18, 756-761.

[7] Singh, P., Anand, S., 2018. Improved performance pulse triggered flip-flop. International Journal of Applied Engineering Research. 13(8), 93-96.

[8] Phyu, M.W., Goh, W.L., Yeo, K.S., 2006. Low-power/high-performance explicit-pulsed flip-flop using static latch and dynamic pulse generator. IEEE Proceedings-Circuits, Devices and Systems. 153(3), 253-260.

[9] Ameena, S., Yasmeen, W., 2020. A low power explicit pulse-triggered D flip-flop based on data feed-through scheme. Solid State Technology. 63(2s), 4646-4657.

[10] John, K., Kumar, R.V., Kumar, S.S., 2021. Low power pulsed flip-flop with clock gating and conditional pulse enhancement. International Journal of Automation and Control. 15(2), 259-274.

[11] Lin, J.F., 2014. Low-power pulse-triggered flip-flop design based on a signal feed-through. IEEE Transactions on Very Large Scale Integration (VLSI) Systems. 22(1), 181-185. DOI: https://doi.org/10.1109/tvlsi.2012.2232684

[12] Karimi, A., Rezai, A., Hajhashemkhani, M.M., 2018. A novel design for ultra-low power pulse-triggered D flip-flop with optimized leakage power. Integration. 60, 160-166.

[13] Panahifar, E., Hassanzadeh, A., 2017. A modified signal feed-through pulsed flip-flop for low power applications. International Journal of Electronics and Telecommunications. 63(3), 241-246.

[14] Pan, D., Ma, C., Cheng, L., et al., 2019. A highly efficient conditional feedthrough pulsed flip-flop for high-speed applications. IEEE Transactions on Very Large Scale Integration (VLSI) Systems. 28(1), 243-251.

[15] Prakalya, S.B., 2021. Power efficient clock pulsed D flip flop using transmission gate. Turkish Journal of Computer and Mathematics Education (TURCOMAT). 12(11), 1544-1555.

[16] Consoli, E., Palumbo, G., Pennisi, M., 2011. Reconsidering high-speed design criteria for transmission-gate-based master-slave flip-flops. IEEE Transactions on Very Large Scale Integration (VLSI) Systems. 20(2), 284-295.

[17] Islam, R., Guthaus, M.R., 2015. Low-power clock distribution using a current-pulsed clocked flip-flop. IEEE Transactions on Circuits and Systems I: Regular Papers. 62(4), 1156-1164.

[18] Wu, Q., Pedram, M., Wu, X., 2000. Clock-gating and its application to low power design of sequential circuits. IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications. 47(3), 415-420.

[19] Srinivasan, N., Prakash, N.S., Shalakha, D.,

et al., 2015. Power reduction by clock gating technique. Procedia Technology. 21, 631-635.

[20] John, K., RS, V.K., Kumar, S.S., 2019. Effect of clock gating in conditional pulse enhancement flip-flop for low power applications. Indonesian Journal of Electrical Engineering and Informatics (IJEEI). 7(2), 357-365.

[21] Sharma, D.K., 2012. Effects of different clock gating techniques on design. International Journal of Scientific & Engineering Research. 3(5), 1-4.

[22] Kumar, C.A., Madhavi, B.K., Kishore, K.L., 2021. Enhanced clock gating technique for power optimization in SRAM and sequential circuit. Journal of Automation, Mobile Robotics and Intelligent Systems. 15(2), 32-38.

[23] Shinde, J., Salankar, S.S., 2011. Clock gating—A power optimizing technique for VLSI circuits. 2011 Annual IEEE India Conference; 2011 Dec 16-18. Hyderabad. New York: IEEE. p. 1-4.

[24] Wang, X., Robinson, W.H., 2010. A low-power double edge-triggered flip-flop with transmission gates and clock gating. 2010 53rd IEEE International Midwest Symposium on Circuits and Systems; 2010 Aug 01-04. Seattle. New York: IEEE. p. 205-208.

[25] Pal, S., Gupta, V., Islam, A., 2021. Variation resilient low-power memristor-based synchronous flip-flops: Design and analysis. Microsystem Technologies. 27(2), 525-538.

BILINGUAL PUBLISHING GROUP

Journal of Electronic & Information Systems
https://journals.bilpubgroup.com/index.php/jeis

ARTICLE

Underwater Image Enhancement Using MIRNet

***M.Thamarai*[1*], *S P Aruna*[2], *Kamesh Sonti*[1], *P.Sudheer Chakravarthi*[1]**

[1] *Sri Vasavi Engineering College, Pedatadepalli, Andhra Pradesh, 534101, India*
[2] *Skilltroniks Technologies, Tadepalliudem, Andhra Pradesh, 534101, India*

ABSTRACT

In recent years, enhancement of underwater images is a challenging task, which is gaining priority since the human eye cannot perceive images under water. The significant details underwater are not clearly captured using the conventional image acquisition techniques, and also they are expensive. Hence, the quality of the image processing algorithms can be enhanced in the absence of costly and reliable acquisition techniques. Traditional algorithms have certain limitations in the case of these images with varying degrees of fuzziness and color deviation. In the proposed model, the authors used a deep learning model for underwater image enhancement. First, the original image is pre-processed by the white balance algorithm for colour correction and the contrast of the image is improved using the contrast enhancement technique. Next, the pre-processed image is given to the MIRNet for enhancement. MIRNet is a deep learning framework that can be used to enhance the low-light level images. The enhanced image quality is measured using peak signal-to-noise ratio (PSNR), root mean square error (RMSE), and structural similarity index (SSIM) parameters.

Keywords: Underwater; Deep learning; MIRNet; Peak signal-to-noise ratio; Structural similarity index

1. Introduction

Image processing can be used to perform some operations on an image to extract some useful information from it. It is one branch of signal processing where the input is a 2-D signal (image) and the output may be an image or an attribute associated with it. Nowadays, image processing is growing rapidly in the core research area within engineering, medicine and other disciplines too [1].

In image processing, underwater image enhancement plays a crucial role and vision applications over the past few years. The images taken underwater

*CORRESPONDING AUTHOR:
M.Thamarai, Sri Vasavi Engineering College, Pedatadepalli, Andhra Pradesh, 534101, India; Email: mthamarai2014@gmail.com

ARTICLE INFO
Received: 26 March 2023 | Revised: 19 April 2023 | Accepted: 20 April 2023 | Published Online: 24 April 2023
DOI: https://doi.org/10.30564/jeis.v5i1.5600

CITATION
Thamarai, M., Aruna, S.P., Sonti, K., et al., 2023. Underwater Image Enhancement Using MIRNet. Journal of Electronic & Information Systems. 5(1): 36-44. DOI: https://doi.org/10.30564/jeis.v5i1.5600

are affected by various lighting and environmental conditions; hence the quality of the image is degraded. The underwater image suffers from degradation due to scattering and absorption. The scattering and absorption process of light in water influences the overall performance of the systems underwater [2]. Forward scattering leads to blurring of the image features, and backward scattering limits the contrast of the image. Similar is the color fading issue, whereby colors like red and yellow almost disappear with increasing depths, which is the reason for the domination of either the blue or the green color. The underwater images are specified by their poor visibility since light is exponentially attenuated as it travels in water and the scenes result poorly contrasted and hazy as shown in **Figure 1(a-c)**. Hence, it is necessary to enhance the underwater images for analyzing its quality, and to prepare the image for further processing [3].

(a) Underwater Fish image

(b) Coral reef image

(c) Under water image with light scattering

Figure 1. Sample underwater images.

The rest of the paper is organized as follows: Section 2 reviews the literature on image processing underwater. Section 3 presents a new method for enhancing the quality of underwater image. Section 4 discusses the simulation results obtained by using our model and comprehensive analysis of the model by evaluating various metrics. Finally, Section 5 describes the conclusion of the work.

2. Literature survey

Schettini et al. [4] review the enhancement and restoration methods for underwater image processing. They discussed light propagation in water, image color correction, lightning problems, and various quality assessment models.

Boudhane et al. [5] proposed a method for pre-processing and fish localization in underwater images by using a mean-shift algorithm for image segmentation and the Poisson-Gauss mixture algorithm for noise reduction, and tested their model under different underwater conditions.

Ancuti et al. [2] performed a fusion of two images (color compensated and white balance version) and then transforms the edges and color contrast to the output images.

Daway et al. [6] performed underwater image enhancement by changing the color content in the image from RGB to YCbCr space. They used Rayleigh distribution along with an integrated color model and calculated no-reference-image quality metrics.

Li et al. [7] created an underwater image enhancement benchmark with 950 raw images, 890 reference images, and 60 challenging images. They also provided an underwater image enhancement network named Water-Net and made the dataset public.

Han et al. [8] proposed a convolution neural network (CNN) based method by combining the max-RGB method and the shades of grey method for detecting the underwater objects.

Wang et al. [9] proposed an underwater image enhancement CNN using two color spaces that integrate RGB color space and HSV color space and evaluate their method with qualitative and quantitative comparisons on both synthetic and raw images.

Zheng et al. [10] proposed a CNN-based network for enhancing the underwater images using an end-to-end defogging module. They also added a cross-layer connection, and pooling pyramid module to improve the defogging network's ability to extract the required information.

3. Proposed work

This section discusses the proposed methodology for enhancing the quality of underwater images.

3.1 Block diagram

The proposed enhancement process of the model is shown in **Figure 2**. The first step is the acquisition of RGB images from the dataset. The input image is pre-processed by the White-Balance algorithm for colour correction and the contrast of the image is improved using the contrast enhancement technique. White-Balance aims at improving the image aspect, primarily by removing the undesired color castings due to various illumination or medium attenuation properties [11]. White-Balance discusses the visible color white under specific lighting conditions affects the hue of all other colors.

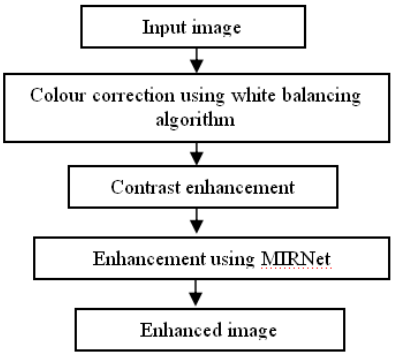


Figure 2. Proposed enhancement process.

White balance algorithm

White-Balance process aims at improving the image aspect, primarily by removing the undesired color castings due to various illumination or medium attenuation properties. White-Balance defines what the color white looks like in specific lighting conditions, which also affects the hue of all other colors. Therefore, when the White-Balance is off, the digital photos and recordings may appear to have a certain hue cast over the image. In underwater the perception of color is highly correlated with the depth, and an important problem is the green-bluish appearance that needs to be rectified. As the light penetrates the water, the attenuation process affects selectively the wavelength spectrum, thus affecting the intensity and the appearance of a colour surface. Since the scattering attenuates more the long wavelengths than the short ones, the color perception is affected down in deeper water. In practice, the attenuation and the loss of color also depend on the total distance between the observer and the scene. Despite white balance being crucial to recover the color, using this correction step is not sufficient to solve the dehazing problem since the edges and details of the scene have been affected by the scattering.

The White-Balance algorithm has three types. They are the White Patch algorithm, Gray World algorithm, and Ground Truth algorithm. From these three algorithms for selecting the best algorithm, we applied the histogram plots on output images of three methods. We found that the Ground Truth algorithm is the best one as compared to the remaining algorithms.

Ground truth algorithm for white-balance

Ground Truth is a term used in statistics and machine learning that means checking the results of machine learning for accuracy against the real world. The term is borrowed from meteorology, where "Ground Truth" refers to information obtained onsite. The term implies a kind of reality checks for machine learning algorithms. The Ground Truth of a satellite image means the collection of information at a particular location. It allows satellite image data to be related to real features and materials on the ground. This information is frequently used for the calibration of remote sensing data and compares the result with Ground Truth. So far, we have made assumptions about how the color spaces behave in our images. Instead of making assumptions for enhancing our images, we select a patch (portion of an image) and use that patch to recreate our desired image. Having selected the patch, we proceed to enhance

our image. For this purpose, we can do it two ways: 1) MAX method—normalize each channel of the original image to the maximum value of each channel of the region, 2) MEAN method—normalize each channel of the original image to the mean value of each channel of the region. The output is slightly closer to the white patch output but the latter is brighter. It also emphasized the color of the lily, but instead of highlighting the color of the pads, it only brightened it. For the Ground Truth algorithm, the output image depends greatly on the choice of the patch image. So, the patch is chosen wisely by visualizing the enhanced image based on the type of application.

Next, the colour corrected image can be passed through contrast enhancement. This step aims to increase image perception by the human eye. This technique plays a major role to bring out the existing information within the low dynamic range of that grey level image [12]. It is required to perform the operations like contrast enhancement and reduction or removal of noise to improve the image quality. Adaptive Histogram Equalization is used for contrast enhancement of the image. Adaptive Histogram Equalization is different from Histogram Equalization in that it computes multiple histograms for each individual part of the image and uses them to spread the image's brightness levels. As a result, it is appropriate for enhancing local contrast in images.

Next, the pre-processed image can be passed through the MIRNet for enhancement [13]. MIRNet is a deep learning framework which can be used to enhance the given image at a low light level.

3.2 MIRNet

It is a feature extraction model which calculates a set of features across various spatial scales and maintains the original high-resolution features for preserving the spatial details. In this process, the features across various resolutions are fused together and repeat this mechanism for representation learning. Also, it is a modern approach to fusing the multi-scale features with the help of a selective kernel network which combines variable receptive fields and faithfully preserves the original features at each spatial resolution. This recursive residual design progressively breaks down the input signal to simplify the overall learning process and allows the construction of a deep neural network as shown in **Figure 3**.

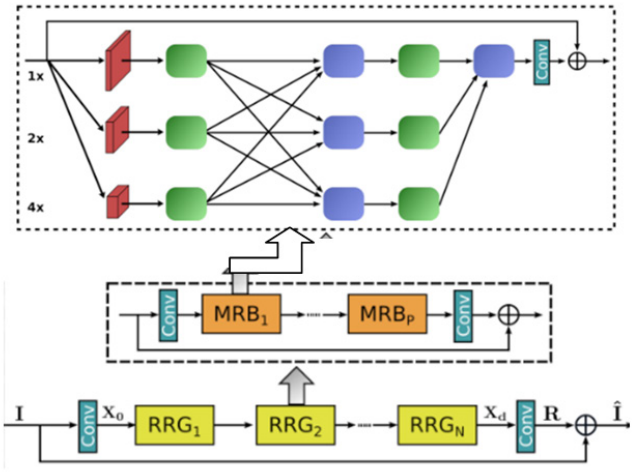


Figure 3. MIRNet architecture.

Selective kernel feature fusion (SKFF)

The SKFF module can perform the dynamic adjustment of receptive fields via i) **Fuse** and ii) **Select** operations. The first operator will generate the global feature descriptors by summing the information from multi-resolution streams. The second operator uses the descriptors for feature maps recalibration followed by aggregation as shown in **Figure 4**.

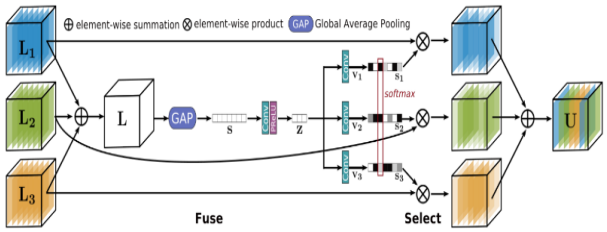


Figure 4. Structure of SKFF.

Dual attention unit (DAU)

The DAU can extract the features from the convolution streams. While the previous block fuses the information across multi-resolution branches, we also need some mechanism to share information within a feature tensor, both along the spatial and the channel dimensions which are done by the DAU block. The DAU reduce less useful features and only allows more informative features to pass further. This process of feature recalibration can be achieved

by two mechanisms: **Channel** and **Spatial Attentions** as shown in **Figure 5**.

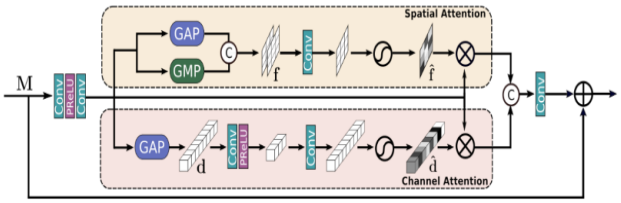
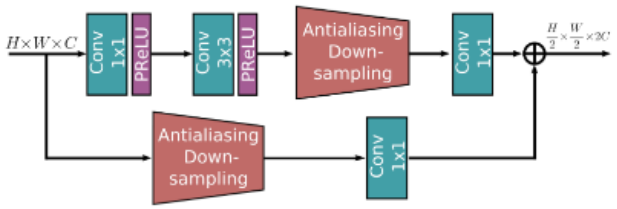


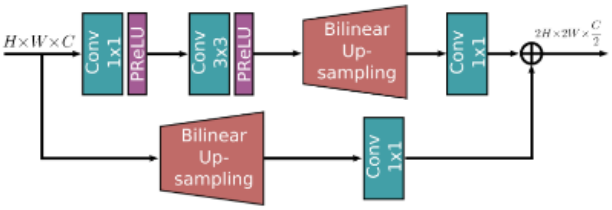
Figure 5. Structure of DAU.

Multi scale residual block (MRB)

The MRB can receive rich contextual information from low-resolutions and generate a spatially-precise output by maintaining high-resolution representations. It consists of multiple (three in this paper) fully-convolution streams which are parallel connected. The MIRNet employs a recursive residual design (with skip connections) to ease the flow of information during the learning process. To maintain the residual nature of our architecture, down sampling and up sampling operations are performed between residual resizing modules as shown in **Figure 6(a-b)**.



(a) Down sampling module in MRB



(b) Up sampling module in MRB

Figure 6. Structure of MRB.

3.3 Software requirements

The software used in our implementation is Python. Python is an object-oriented, high-level language, interpreted, dynamic and multipurpose programming language [14]. The libraries include:

- Numpy
- Keras
- Matplotlib
- Scikit-learn

4. Simulation results

In this section, the simulation results carried out in our work are presented which follows the comparison of results.

4.1 Dataset used

The dataset used in our implementation is the underwater image enhancement benchmark (UIEB) dataset [7], which includes 890 raw underwater images and corresponding reference images.

4.2 Performance metrics

The performance metrics used in our implementation include peak signal-to-noise ratio (PSNR), mean square error (MSE), root mean square error (RMSE), and structural similarity index (SSIM) which are defined as:

$$PSNR = 10.\log_{10}\left(\frac{255^2}{MSE}\right)$$

$$MSE = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left[I(i,j) - I'(i,j)\right]^2$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}$$

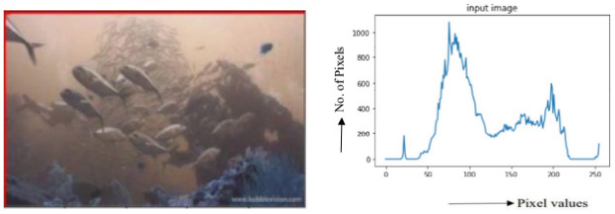
$$SSIM(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_1^2 + \mu_2^2 + c_1)(\sigma_1^2 + \sigma_2^2 + c_2)}$$

4.3 Simulation results

From the UIEB dataset, different raw images are taken as inputs and applied the proposed algorithm and obtained the white balanced image, contrast enhanced image, MIRNet output image, and corresponding histograms for each image. In this paper, the input and output images and corresponding histogram plots for four different raw images are present-

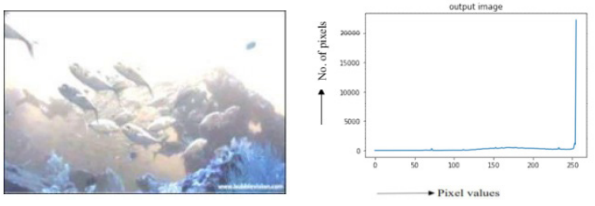
ed and are shown in **Figures 7-10** respectively.

❖ **Raw image-1:** Image size 208 × 319



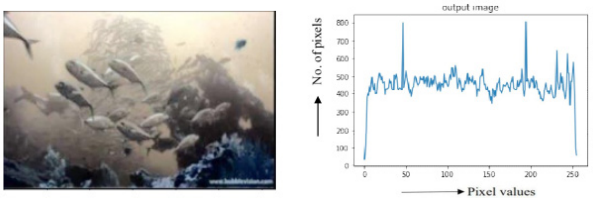
(i) (ii)

Figure 7(a). (i) Input image-1 and (ii) histogram of input image-1.



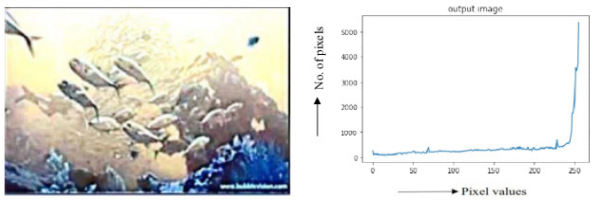
(i) (ii)

Figure 7(b). (i) White balanced output for image-1 and (ii) histogram of white balanced output image-1.



(i) (ii)

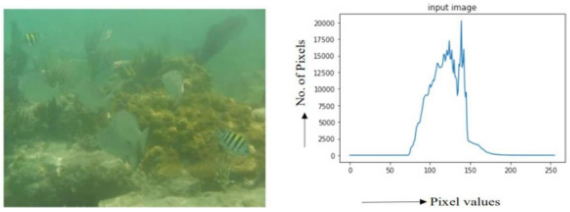
Figure 7(c). (i) Contrast enhanced output for image-1 and (ii) histogram of contrast enhanced output image-1.



(i) (ii)

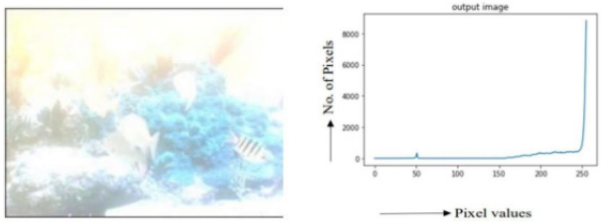
Figure 7(d). (i) Enhanced output for image-1 using MIRNet and (ii) histogram of enhanced output image-1.

❖ **Raw image-2:** Image size 768 × 1024



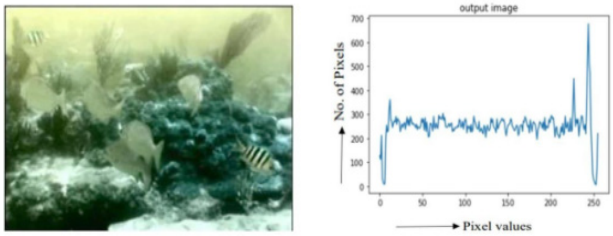
(i) (ii)

Figure 8(a). (i) Input image-2 and (ii) histogram of input image-2.



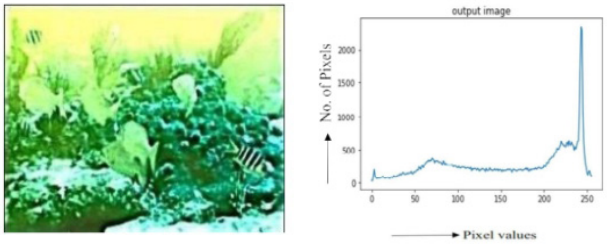
(i) (ii)

Figure 8(b). (i) White balanced output for image-2 and (ii) histogram of white balanced output for image-2.



(i) (ii)

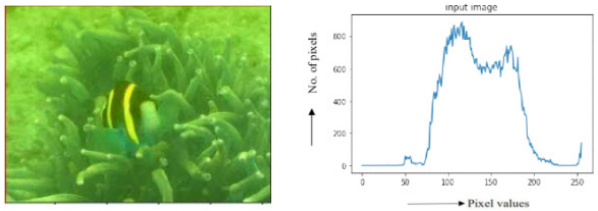
Figure 8(c). (i) Contrast enhanced output for image-2 and (ii) histogram of contrast enhanced output image-2.



(i) (ii)

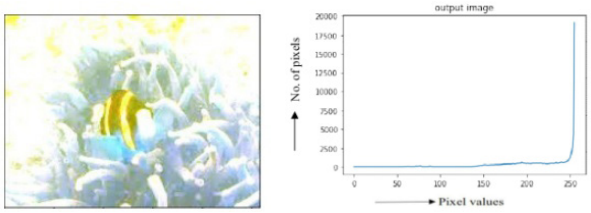
Figure 8(d). (i) Enhanced output for image-2 and (ii) histogram of enhanced output image-2.

❖ **Raw image-3:** Image size 237 × 315



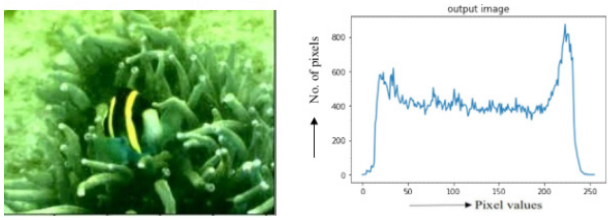
(i) (ii)

Figure 9(a). (i) Input image-3 and (ii) histogram of input image-3.



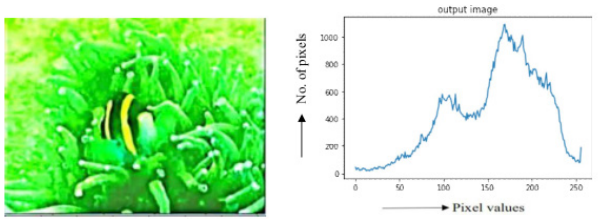
(i) (ii)

Figure 9(b). (i) White balanced output for image-3 and (ii) histogram of white balanced output image-3.



(i) (ii)

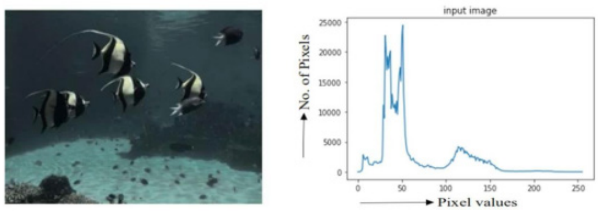
Figure 9(c). (i) Contrast enhanced output for image-3 and (ii) histogram contrast enhanced output image-3.



(i) (ii)

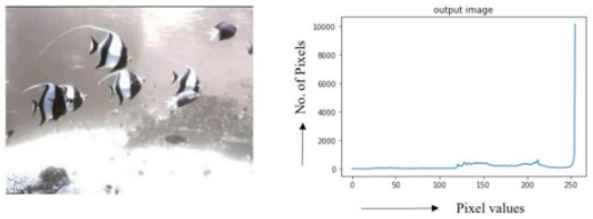
Figure 9(d). (i) Enhanced output for image-3 and (ii) its histogram.

❖ **Raw image-4:** Image size 683 × 910



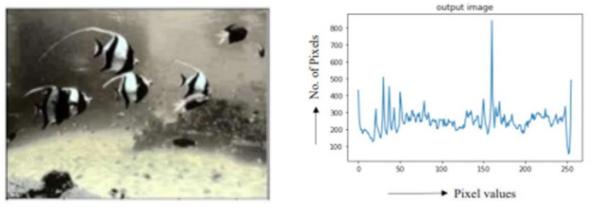
(i) (ii)

Figure 10(a). (i) Input image-3 and (ii) histogram of input image-3.



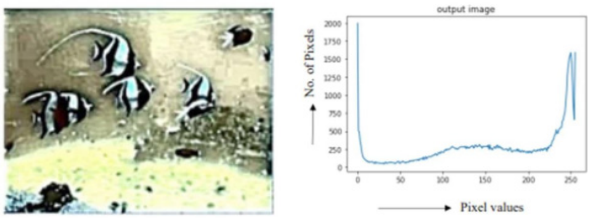
(i) (ii)

Figure 10(b). (i) White balanced output for image-4 and (ii) histogram of white balanced output image-4.



(i) (ii)

Figure 10(c). (i) Contrast enhanced output for image-4 and (ii) histogram of enhanced image-4.



(i) (ii)

Figure 10(d). (i) Enhanced output for image-4 and (ii) histogram of enhanced image-4.

The performance metrics for the above mentioned four images ere calculated and compared with the metrics for the color corrected image and the output enhanced image. The evaluated metrics for the images are represented in **Tables 1-4**.

Table 1. Result comparison for image-1.

Metric	Color corrected image using white balance	Enhanced image using MIRNet
RMSE	73.9	34.7
PSNR	10.7	17.3
SSIM	(0.4, 0.4)	(0.8, 0.8)

Table 2. Result comparison for image-2.

Metric	Color corrected image using white balance	Enhanced image using MIRNet
RMSE	77.3	26.5
PSNR	10.3	19.6
SSIM	(0.3,0.4)	(0.8,0.8)

Table 3. Result comparison for image-3.

Metric	Color corrected image using white balance	Enhanced image using MIRNet
RMSE	53.5	24.5
PSNR	13.5	20.3
SSIM	(0.4, 0.4)	(0.7, 0.7)

Table 4. Result comparison for image-4.

Metric	Color corrected image using white balance	Enhanced image using MIRNet
RMSE	115.8	29.16
PSNR	6.8	18.83
SSIM	(0.2, 0.3)	(0.6, 0.6)

5. Conclusions

A deep learning framework based technique MIR-Net is proposed in our paper for underwater image enhancement which transforms images to provide a better representation of the content present in the image. First, the color of the input image is corrected using a technique called the White-Balance algorithm. Next, the contrast of the color corrected images is improved using an adaptive histogram equalization method. Finally, the pre-trained deep learning model MIRNet is used to enhance the image. The quality of the output images is justified in terms of PSNR and SSIM parameters and tabulated. Hence, we conclude that the proposed method can perform underwater image enhancement for low light, low contrast underwater images effectively without any loss of information.

Future Scope

The proposed method can be extended in the future with suitable pre-processing techniques to enhance the blurred underwater images which result from the light scattering problem of the underwater image capturing system.

Conflict of Interest

There is no conflict of interest.

References

[1] Acharya, T., Ray, A.K., 2005. Image processing: Principles and applications. John Wiley & Sons: New York.

[2] Ancuti, C.O., Ancuti, C., De Vleeschouwer, C., et al., 2017. Color balance and fusion for underwater image enhancement. IEEE Transactions on Image Processing. 27(1), 379-393.

[3] Raveendran, S., Patil, M.D., Birajdar, G.K., 2021. Underwater image enhancement: A comprehensive review, recent trends, challenges and applications. Artificial Intelligence Review. 54(7), 5413-5467.

[4] Schettini, R., Corchs, S., 2010. Underwater image processing: State of the art of restoration and image enhancement methods. EURASIP Journal on Advances in Signal Processing. 2010, 1-14.

[5] Boudhane, M., Nsiri, B., 2016. Underwater image processing method for fish localization and detection in submarine environment. Journal of Visual Communication and Image Representation. 39, 226-238.

[6] Daway, H.G., Daway, E.G., 2019. Underwater image enhancement using colour restoration based on YCbCr colour model. IOP Conference Series: Materials Science and Engineering. 571(1), 012125.

[7] Li, C., Guo, C., Ren, W., et al., 2019. An underwater image enhancement benchmark dataset and beyond. IEEE Transactions on Image Processing. 29, 4376-4389.

[8] Han, F., Yao, J., Zhu, H., et al., 2020. Underwater image processing and object detection based on deep CNN method. Journal of Sensors. 2020(9), 1-20.

[9] Wang, Y., Guo, J., Gao, H., et al., 2021. UIEC^2-Net: CNN-based underwater image enhancement using two color space. Signal Processing: Image Communication. 96, 116250.

[10] Zheng, M., Luo, W., 2022. Underwater image enhancement using improved CNN based de-

fogging. Electronics. 11(1), 150.

[11] Thai, B., Deng, G., Ross, R., 2017. A fast white balance algorithm based on pixel greyness. Signal, Image and Video Processing. 11, 525-532.

[12] Hashemi, S., Kiani, S., Noroozi, N., et al., 2010. An image contrast enhancement method based on genetic algorithm. Pattern Recognition Letters. 31(13), 1816-1824.

[13] Zamir, S.W., Arora, A., Khan, S., et al., 2022. Learning enriched features for fast image restoration and enhancement. IEEE Transactions on Pattern Analysis and Machine Intelligence. 45(2), 1934-1948.

[14] Dey, S., 2018. Hands-on image processing with Python: Expert techniques for advanced image analysis and effective interpretation of image data. Packt Publishing Ltd: Birmingham.

BILINGUAL PUBLISHING GROUP

Journal of Electronic & Information Systems
https://journals.bilpubgroup.com/index.php/jeis

ARTICLE

The Refractive Effect of k-Factor on Radio Propagation over Lokoja, Nigeria

Akinsanmi Akinbolati*, Florence N. Ikechiamaka, Akogwu O. Isaiah

Department of Physics, Federal University Dutsin-Ma, 821101, Nigeria

ABSTRACT

The effective earth radius factor (k-factor) has a refractive propagation effect on transmitted radio signals thus making its study necessary for the proper planning of terrestrial radio links and power budget. This study was carried out over the city of Lokoja, Nigeria, using ten years (2011 to 2020) atmospheric data of temperature, pressure and humidity both at the surface (12 m) and at 100 m AGL. The data were retrieved from European Centre for Medium-Range Weather Forecasts (ECMWF) ERA5. The k-factor yearly variation follows the same trend with minimum and maximum values obtained during dry and wet season months respectively. In addition, the highest mean value of 1.00042 was recorded in the month of August while the lowest value of 1.00040 was recorded in the month of January with an overall mean value of 1.0003. This value is less than the recommended standard of 1.33 by ITU-R. The propagation effect corresponding to $k < 1.33$ is sub-refractive. The implication of this on radio wave propagation, especially terrestrial communications is that transmitted wireless signal is prone to losses. This can be mitigated through an effective power budget: Choice of transmitting antenna's height and gain, so as to improve the Quality of Service over the study area.

Keywords: Effective earth radius factor (k-factor); Refractive effect; Terrestrial radio link; Radio signal; Power budget

1. Introduction

Generally, the dynamics in the atmosphere can lead to attenuation of radio signal propagated from the source which is the transmitter to the destination also known as the receiver [1]. The interaction of propagated radio signals with the path of propagation through various mechanisms often leads to signal degradation [2]. This has made path loss-related studies by radio scientists and engineers imperative in

*CORRESPONDING AUTHOR:
Akinsanmi Akinbolati, Department of Physics, Federal University Dutsin-Ma, 821101, Nigeria; Email: aakinbolati@fudutsinma.edu.ng

ARTICLE INFO
Received: 21 March 2023 | Revised: 20 April 2023 | Accepted: 21 April 2023 | Published Online: 28 April 2023
DOI: https://doi.org/10.30564/jeis.v5i1.5583

CITATION
Akinbolati, A., Ikechiamaka, F.N., Isaiah, A.O., 2023. The Refractive Effect of k-Factor on Radio Propagation over Lokoja, Nigeria. Journal of Electronic & Information Systems. 5(1): 45-50. DOI: https://doi.org/10.30564/jeis.v5i1.5583

radio communication industries [3]. Radio scientists and engineers need knowledge of the effective earth radius factor (k-factor) in a local environment for the proper planning of terrestrial radio links and power budget. If the troposphere is homogenous and turbulent free, any wave propagated into the troposphere parallel to the earth's surface will follow the earth's curvature [4] and be maximally received. However, due to the inhomogeneity of the troposphere, radio wave propagated undergoes bending depending on the tropospheric condition which also depends on the interaction of the radio climatic factors in the environment [5].

The k-factor is useful in the prediction of local radio wave propagation conditions. Typically, a design value of k = 1.33 is often assigned in locations where the values are not known. Studies have also shown that its value is location based and should not be assumed constant for all environments [4]. Since k-factor is weather and climate dependent, it is therefore imperative for regular studies to be carried out using up-to-date data in order to capture the effects of climate change. It also represents a spatial average, which can only otherwise be obtained from simultaneous meteorological surroundings along the propagation path. The k-factor is the radius of a hypothetical spherical Earth, without an atmosphere, for which propagation paths follow straight lines [6].

1.1 Radio refractivity

The radio refractivity N can be expressed as:

$$N=\frac{77.6}{T}\left(P+4810\frac{e}{T}\right)(N\text{-}units) \quad (1)$$

It depends on atmospheric parameters of pressure P (hPa), temperature T (K) and water vapour pressure e (hPa) [7,8].

1.2 k-factor

The k-factor can be obtained using two refractivity values as:

$$\frac{dN}{dh}=\frac{N_2-N_1}{h_2-h_1} \quad (2)$$

where N_1 and N_2 are the refractivity values at heights h_1 and h_2 respectively. In this study h_1 and h_2 are at the surface (12 m) and 100 m Above Ground Level (AGL) respectively. The k-factor can be expressed as indicated in Equations (3)-(4) [4,5].

$$k=\frac{1}{\left(1+a\left(\frac{dN}{dh}\right)\right)} \quad (3)$$

where a is the radius of the earth (a = 6371 km = 344 mni) and dN/dh is the rate of change of refractivity indices with height.

k-factor can also be expressed as indicated in Equation (4), which was used for k-factor computations in this work.

$$k=\frac{1}{1+\frac{\left(\frac{dN}{dh}\right)}{157}} \quad (4)$$

2. Review of related work

Transmission of a radio signal in the lower atmosphere is affected by many processes which include variations in air temperature, pressure, and humidity. These variations in weather parameters often result in refractivity changes [9]. These changes can result in abrupt changes in the propagation direction of a radio signal resulting in signal loss. Based on this premise, studies on the effect of radio climatic factors on propagated signals have become imperative for the planning of a reliable radio link in a region. Adediji and Ajewole [8] studied the vertical refractivity gradient in Akure, Nigeria by measuring atmospheric variables using integrated sensor suits (ISS) at different heights above ground level. Results show that propagation conditions have varying degrees of occurrence. Oyedum et al. [10] worked on reduced sea level refractivity in Minna, Central Nigeria. Ayantunji et al. [11] studied the seasonal and diurnal variation of surface radio refractivity in Akure, Nsuka, Minna, Sokoto and Jos (in Nigeria). The result of the work revealed higher values of surface refractivity during the wet season in Nigeria compared to the dry season months. Propagation of radio signal in the troposphere is affected by the interaction of signal with

some primary (air temperature, pressure, and humidity) and secondary (radio refractivity, refractivity gradient, k-factor, etc.) radio-climatic factors. Many works have been done locally and internationally, but due to climate change it is necessary that studies employ up-to-date data to ensure the high reliability of findings.

A design value of 1.33 is often assigned for k-factor in line of sight link especially where information about the actual value of k-factor for that location is not available [7,4]. P. E. Okpani et al. [12] investigated the effect of radio climatic variables on signal propagation over Nsukka, Nigeria. Results have shown that k-factor values ranged from 1.555-1.652. Ojo et al. [4] worked on the characterization of secondary radio-climatic variables for microwave and millimeter wave link design in Nigeria, using five years of data (2009-2013). The study location included Akure, Enugu, Minna, Jos, and Sokoto cities. Results revealed average values of 1.476, 1.940, 1.860 and 1.287 respectively in Akure, Minna, Jos and Sokoto respectively. Ukhurebor and Odesanya [5] investigated k-factor over Auchi area of Edo State, South-South, Nigeria. The work determined k-factor mean value of 1.470 over Auchi which is slightly greater than ITU-R standard. Abu-Almal and Al-Ansari [6] calculated the k-factor and point refractivity gradient in United Arab Emirates (UAE). Fourteen years of radiosonde meteorological data were employed. Results revealed k-factor monthly variation from 1.43 to 3.17.

3. Methodology

This study was carried out in the city of Lokoja. Lokoja is the state capital of Kogi State, in North Central, Nigeria. Secondary atmospheric data of temperature, pressure and humidity both at the surface (12 m) and at 100 m AGL for Lokoja were retrieved from ECMWF ERA5 in December 2021. The data are high-resolution satellite-data with high reliability covering 2011-2020. ECMWF is an independent intergovernmental organization, which was established in 1975. They produced global numerical weather forecasts for worldwide users [13]. Equations (1)-(4) were used to determine the k-factor values for the years under study. The mean annual values were used for analysis to enhance the reliability of the results.

Data analysis and ITU-R recommendation on k-factor

Necessary analyses such as sorting, calculation of mean values and plotting of graphs were carried out using Excel. The ITU-R Recommendation on k-factor standards on refractive conditions of the atmosphere and the subsequent attenuation effect was equally used at achieving the set objectives. The ITU-R Recommendation on k-factor and associated propagation effects are presented in **Table 1**. The determined values of k-factor obtained for the years and seasons were compared with ITU-R standard in order to predict the refractive propagation effects over the study areas.

Table 1. ITU-R standards on k-factor and the associated propagation effects [5,7,14].

k-factor's range of values	Propagation effect on radio communications
$k = 1.33$ standard atmosphere	In this case, radio signals are transmitted along a straight line part on the earth's surface and go into space unimpeded.
$1.33 > k > 0$ Sub-refraction	Here, a portion of the radio wave (signal) propagates abnormally away from the earth surface, resulting to interference and coverage limitation.
$\infty > k > 1.33$ Super-refraction	Here, the radio wave signals (e. g microware link, GSM, satellite) spread irregularly toward the earth's surface thus, surface extending the radio horizon and merge the path clearance giving rise to irregular huge waves above the line of view due to multiple reflection.
$-\infty < k < 0$ Ducting	Here, there will be ducting which will make the radio waves to bend downwards with a curvature greater than the earth's.

4. Results

Figure 1 presents the yearly variation of k-factor covering 2011 to 2020 over Lokoja. The variation follows the same trend for the ten years under study with minimum and maximum values obtained during dry and wet season months respectively. **Figure 2** presents the mean variation of k-factor over the

months of the years under study using mean values for the ten years in Lokoja. Observation from the figures shows lower values were obtained during the dry season months of January-March and October-December while higher values were obtained during the wet season months of April to September. **Tables 2** and **3** present the determined correlation coefficient(s) between k-factor and atmospheric parameters at the surface (12 m) and at 100 m respectively.

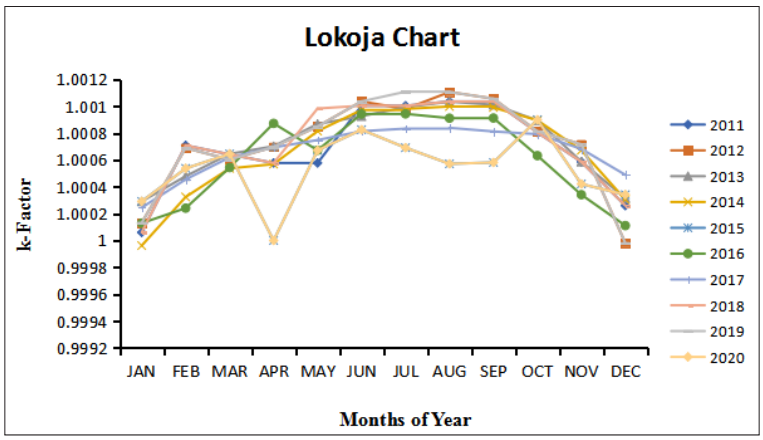


Figure 1. Yearly variation of k-factor covering the ten years (2011-2020) under study over Lokoja.

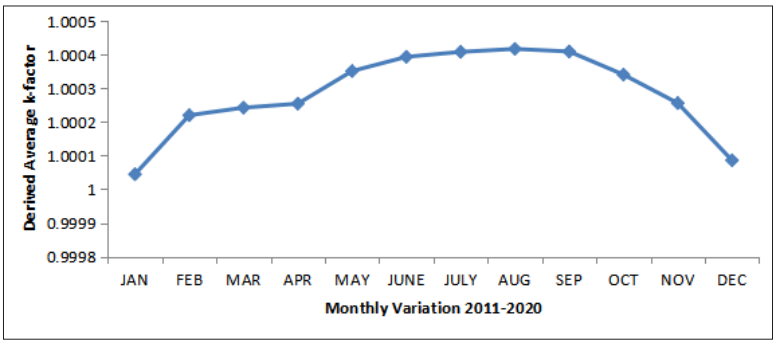


Figure 2. The mean monthly variation of k-factor for the ten years under study over Lokoja.

Table 2. Correlation coefficient between k-factor and atmospheric parameters at surface (12 m, AGL).

	Temp (C)	Humidity (%R.H)	Press. (hPa)	Rainfall (mm)	k-factor
Temp	1				
Humidity	–0.810	1			
Press.	–0.923	0.687	1		
Rainfall	–0.603	0.849	0.578	1	
k-factor	–0.620	0.949	0.715	0.784	1

Table 3. Correlation coefficient between k-factor and atmospheric parameters at 100 m, AGL.

	Temp (C)	Humidity (%R.H)	Press. (hPa)	Ammount Rainfall (mm)	k-factor
Temp	1				
Humidity	–0.846	1			
Press.	0.923	–0.676	1		
Rainfall	–0.756	0.840	–0.590	1	
k-factor	–0.872	0.955	–0.724	0.784	1

5. Discussion

In addition, the highest value of 1.00042 was recorded in the month of August while the lowest value of 1.0004 was recorded in the month of January. The overall mean value for the ten years under study is 1.00029 or approximately equal to 1.0003. This value is less than the recommended standard of 1.33 (i.e $1.33 > k > 0$) standard by the International Telecommunication Union, Radio Study Group (ITU-Rec. 530-16, 2015). The propagation effect corresponding to $k < 1.33$ is sub-refraction; by interpretation, the atmosphere for the ten years under study in Lokoja is sub-refractive. The effect of this on radio wave propagation is that a certain portion of the propagated radio signals has the tendency to propagate abnormally by not following the earth's curvature. This often leads to interference and signal degradation before the expected points of reception. In this kind of situation the quality of reception of radio signals on the microwave, Satellite, VHF (Very High Frequency) and UHF (Ultra High Frequency) bands are affected. As a result of the finding of this work which employed ten years of recent data, the probability that Lokoja will experience sub-refraction in the subsequent years ahead is high. The implication of this on radio wave propagation especially terrestrial communications is that transmitted wireless signal is prone to losses due to sub-refractive. Based on these results, terrestrial radio communication engineers and various stakeholders of radio propagation systems in Lokoja should make deliberate efforts at mitigating sub-refractive losses. This can be achieved through an effective power budget; choice of transmitter power and antenna height and gain. In addition, existing terrestrial radio communication stations over the study areas may need to increase their antenna's height or by employing the use of repeater station(s), so as to improve the Quality of Service (QoS) over the study location.

The correlation coefficient between k-factor and temperature, pressure, humidity and rainfall at the surface are –0.62, 0.72, 0.95 and 0.79 respectively. Here, a high positive correlation coefficient was obtained between k-factor and humidity and k-factor and rainfall. The implication is that the higher the humidity and rainfall over the location, the higher the k-factor. This is consistent with the results depicted in **Figure 2**. On the other hand, using the atmospheric parameters at 100 m AGL, the correlation coefficient between k-factor and; temperature, pressure, humidity are –0.87, –0.72, 0.96 respectively. Here, a high positive correlation coefficient was obtained between k-factor and humidity again.

6. Conclusions

The variation of the effective earth radius factor (k-factor) for a period of ten years over the city of Lokoja was investigated. The refractive propagation effect of this secondary radio climatic factor on radio communication was revealed as sub-refractive based on the ITU-R standard of categorization. The implication of this on radio wave propagation especially terrestrial communications is that transmitted wireless signal is prone to losses due to sub-refractive. Based on these results, terrestrial radio communication engineers and various stakeholders of radio propagation systems in Lokoja should make efforts at mitigating sub-refractive losses. This can be achieved through an effective power budget; choice of transmitter power and antenna height and gain. In addition, existing terrestrial radio communication stations over the study areas may need to increase their antenna's height or by employing the use of repeater station(s), so as to improve the Quality of Service (QoS) over the study location.

Author Contributions

1) Author 'AA' (#1): Designed the study, made data available, read through the manuscript and made useful corrections.

2) Author 'FNI' (#2): Proof read the manuscript and made a contribution.

3) Author 'AOI' (#3): Carried out the study, wrote the first draft.

Conflict of Interest

There is no conflict of interest.

References

[1] Akinbolati, A., Akinsanmi, O., Ekundayo, K.R., 2016. Signal strength variation and propagation profiles of UHF radio wave channel in Ondo state, Nigeria. International Journal of Microwave and Wireless Technologies. 6(4), 12-27. DOI: https://doi.org/10.5815/ijwmt.2016.04.02

[2] Akinbolati, A., Agunbiade, O.J., 2020. Assessment of error bounds for path loss prediction models for TV white space usage in Ekiti State, Nigeria. International Journal of Information Engineering and Electronic Business. 10(3), 28. DOI: https://doi.org/10.5815/ijieeb.2020.03.04

[3] Akinbolati, A., Ajewole, M.O., 2020. Investigation of path loss and modeling for digital terrestrial television over Nigeria. Heliyon. 6(6), e04101. DOI: https://doi.org/10.1016/j.heliyon.2020.e04101

[4] Ojo, J.S., Ojo, O.L., Akinyemi, P., 2019. Characterization of secondary radioclimatic variables for microwave and millimeter wave link design in Nigeria. Indian Journal of Radio & Space Physics (IJRSP). 46(3), 83-90.

[5] Ukhurebor, K.E., Odesanya, I., 2019. Relationship between meteorological variables and effective earth radius factor over Auchi, Edo State, South-South, Nigeria. Covenant Journal of Physical and Life Sciences (Special Edition). 7(1).

[6] Abu-Almal, A., Al-Ansari, K., 2010. Calculation of effective earth radius and point refractivity gradient in UAE. International Journal of Antennas and Propagation. 245070.

[7] International Telecommunications Union, Radio Study Group (ITU-R), *Recommendation on Radio refractivity,* ITU-Rec. P. 453(2003-2019). https://www.itu.org/rec.p.453

[8] Adediji, A.T., Ajewole, M., 2008. Vertical profile of radio refractivity gradient in Akure South-West Nigeria. Progress in Electromagnetics Research C. 4, 157-168.

[9] Okoro, O.N., Agbo, G.A., 2012. The effect of variation of meteorological parameters on the tropospheric radio refractivity for Minna. Global Journal of Science Frontier Research Physics & Space Science. 12(2).

[10] Oyedum, O.D., Igwe, K.C., Eiche, J.O., et al., 2009. Reduced-to-sea-level refractivity in Minna, Central Nigeria. Natural and Applied Science Journal. 11(2), 1-5.

[11] Ayantunji, B.G., Okeke, P.N., Urama, J.O., 2011. Diurnal and seasonal variation of surface refractivity over Nigeria. Progress in Electromagnetics Research B. 30, 201-222.

[12] Okpani, P.E., Nwofe, P.A., Chukwu, N.O., 2015. Effect of secondary radioclimatic variables on signal propagation in Nsuka, Nigeria. International Research Journal of Natural Sciences. 3(2), 9-17.

[13] The European Centre For Medium-Range Weather Forecasts (ECMWF) ERA5, 2021. Available from: https://www.ecmwf.int/en/forecasts/dataset/ecmwf-reanalysis-v5

[14] Ajayi, G.O., 1989. Physics of the tropospheric radio propagation. Proceeding of the ICTP College on Theoretical and Experimental Radio Propagation Physics; 1989 Feb 6-24; Trieste, Italy. p. 54.

BILINGUAL PUBLISHING GROUP

Journal of Electronic & Information Systems
https://journals.bilpubgroup.com/index.php/jeis

ARTICLE

Introduction to Thermo-Photo-Electronics

Stanislav Ordin

The Russian Academy of Sciences, Saint Petersburg, 119991, Russia

ABSTRACT

Building the foundations of Thermo-Photo-Electronics became possible only after the correction of thermodynamic errors in the traditional theory of semiconductor Electronics and Photo-Electronics. It is these errors that determined the output of the asymptotics of the operating parameters of semiconductor electronic devices, in particular, both the saturation of the limiting clock frequency of processors, and the saturation of the efficiency of both thermoelectric and photoelectric converters. But in semiconductors, although these thermodynamic errors manifested themselves not only in the instrumental, but also in the technological aspect, they could not prohibit semiconductor Electronics itself, unlike Electronics based on other materials. It's just that a number of qualitative mistakes were made in the theory of semiconductor devices and photo devices. In this work, it is shown that the energy band diagram of semiconductor contacts itself was constructed with a significant omission—without taking into account the temperature force on the contact. At the same time, because of the incorrect calculation of currents according to the outdated formulas of Richardson-Langmuir-Deshman, there were also PROHIBITIONS. So the practitioners compensated for the errors of the theory with "empirical corrections". So electronics engineers often made devices not according to a strict theory (which simply did not exist until now), but on a hunch and according to empirical local laws. And only the correction of the historical mistakes made it possible to expand the phenomenology of the description of processes in a Solid Body, on the basis of which it is possible to make calculations of highly efficient elements of Photo-Thermo-Electronics.

Keywords: Phenomenology; Potential barriers; P-n-junction; Prigogine local entropy production; Richardson-Langmuir models; Local thermo-EMF

*CORRESPONDING AUTHOR:
Stanislav Ordin, The Russian Academy of Sciences, Saint Petersburg, 119991, Russia; Email: stas_ordin@mail.ru

ARTICLE INFO
Received: 21 March 2023 | Revised: 18 April 2023 | Accepted: 20 April 2023 | Published Online: 11 May 2023
DOI: https://doi.org/10.30564/jeis.v5i1.5580

CITATION
Ordin, S., 2023. Introduction to Thermo-Photo-Electronics. Journal of Electronic & Information Systems. 5(1): 51-66. DOI: https://doi.org/10.30564/jeis.v5i1.5580

1. Introduction

Ilya Prigogine's discovery of the production of local entropy was not taken into account by Electronics, although it directly relates to micro- and nano-scales. And this phenomenological error in the theory of semiconductor devices is decisive. But it was aggravated by the fact that ALL solid-state physics was "stuck" on the use of flat electron orbitals, which Pauling introduced for "two-dimensional" graphite and for which he received the Nobel Prize. But Pauling himself, as an honest scientist, having discovered his mistake, tried to correct it by introducing "curved" orbitals for graphite itself [1]. Now, after restoring the Planck-Einstein quantization, it is shown how to get the correct orbitals instead of the mystical Schrödinger wave functions [2-4]. And this is the next stage of the rigorous expansion of the phenomenology of Electronics on the basis of the obtained foundations to the phenomenology of Photo-Thermo-Electronics and to the construction of a rigorous theory of semiconductor devices.

The dimensional thermoelectric effect in silicon carbide crystals, which I initially discovered back in the 1980s in the study of contact effects [5-7], has already shown insufficient completeness of thermoelectric phenomenology to describe thermo-EMF [8-10] and output resistance in microstructures with potential barriers (**Figure 1**), which, in fact, prompted me to start a long-term cycle of research on Local (NANO) Effects (formerly called midi-effects).

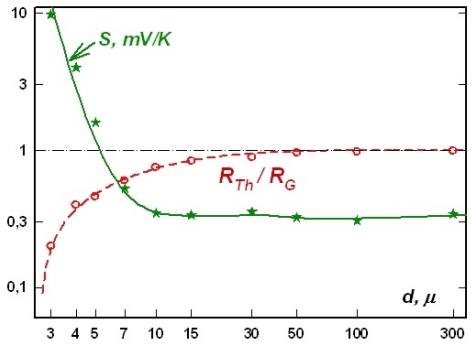


Figure 1. Dependences of thermo-EMF (originally it was assumed that this is the diffuse Seebeck coefficient) and the ratio of the output thermoelectric resistance to the galvanic resistance of a SiC semiconductor crystal depending on its thickness (size along the heat flow).

But a rigorous theoretical extension of the purely diffuse (as analysis showed) theory of thermoelectricity was made me later [11-14], when the existence of Local thermoelectric power was confirmed for a whole class of wide-gap semiconductors [15] and it became clear that not only phenomenology of thermoelectricity, but also the entire phenomenology of the *p-n* junction [16].

And then, neither the laboratory of thermoelectricity of the Ioffe Institute of the USSR Academy of Sciences, nor the International Thermoelectric Society dared to deviate from traditional thermoelectric concepts and, with the need to correctly take into account the concentration force. Take it into account in the same way as it has long been done in the theory of *p-n* junction. And the management of the Global Thermoelectric company, which arranged a radio conference with me at the beginning of this century, summarized both our discussion and the general position of the thermoelectric community by saying: "We are now the world's largest manufacturer of thermoelectric generators based on well-established traditional thermoelectric technology. And where will we be if we switch to the Intel technology you propose for thermoelectricity".

But even at the beginning of the century, even if then at a qualitative level, I already felt that there is a "reverse side of the coin", that phenomenology, and, as a result, calculations of the characteristics of the p-n junction are also not complete without taking into account the temperature force traditionally used in thermoelectricity. I understood and showed at a qualitative level to the management of Intel that Moore's law does not take into account that it was precise because of the lack of taking into account the temperature force that they reached the saturation of the processor clock frequency with increasing miniaturization of individual elements (an increase in the number of elements in a chip).

But Intel also followed a purely empirical path according to its own, mastered technology - first they blind the processor, and then examine what they did. He took so long to respond that I wrote to them: "Dear Sirs, I am not sure that I will live to see the

next millennium. So please decide on my project." And INTEL answered: "Thank you for your patience and humor" and was also afraid to move away from its traditional processor manufacturing technology, and tried to solve the problem incorrectly, but to get around it due to multi-core. But this method of increasing the number of elements, even then it was obvious, would not allow going beyond the logarithmic dependence of the speed of processors on the number of elements in them.

My project "NANO-thermopowers" won the Samsung competition, but our Noble Zhores Alferov took the money from it for his heterostructures, which, as it has now become clear, he did WRONG. So, only local, my co-authors, professors Toru Miyakawa, Wang Nang Wang and Satoru Yamoguchi, helped me in conducting research on Local Effects. And now those early qualitative results of mine have received rigorous experimental and theoretical confirmation and have shown the need to expand not only thermoelectricity, but all Electronics to Thermo-Photo-Electronics.

2. Electronics stagnation analysis

2.1 Reasons for the stagnation of thermoelectricity

Thermoelectric instruments and devices have been actively used since the middle of the last century. But already at the end of the last century, both by the demand for them and by the achievements of their marginal efficiency, it became clear that after reaching a certain level, no further progress was observed. On the contrary, refrigerators began to return from thermoelectric to traditional, mechanical ones, and thermogenerators began to be used only if there is a source of waste heat. And thermoelectricity itself as a science actually degenerated into Materials Science [17].

But historically, it was thermoelectricity that allowed Onsager, based on the Curie theorem, to formulate his Principle of Symmetry of Kinetic Coefficients [8]. But there was no feedback. The primitive theory of Thermoelectricity has been built up to now with the violation of the Onsager Principle, admitted by an amateur physicist, watchmaker Peltier. Moreover, she "passed" by Oleg Losev's discovery of a p-n junction, and past the discovery of "anomalous thermo-EMFs" in it, and past Ilya Prigozhin's discovery of local entropy production. So it is not surprising that in the period when all the bureaucratic science degenerated into the knowledge industry, thermoelectricity, as science, simply ceased to claim scientific Fundamentality, and turned into a factory laboratory serving production with obsolete technologies compared to electronics. That is why, in fact, there is no scientific thermoelectric journal, but only factory reports in the proceedings of the international thermoelectric conference, in which my work was published only for decoration [18]. That is why thermoelectricity was not even presented in international multidisciplinary scientific journals either, until the International Journal of Frontier Studies asked me to send my biography. To which I replied that I am no longer a young man and I can tell a lot about myself. But let me tell you the tragic History of Thermoelectricity. And in response, I received an offer to publish this story of mine "Anomalous thermo-EMF is Local thermo-EMF" [19]. And they published, free of charge, like all 70 of my scientific articles in the Open Access in recent years. So even the development of the fundamental aspects of thermoelectricity—contact phenomena, which I was instructed to study 40 years ago by the last Coryphaeus of Thermoelectricity, Lazar Solomonovich Stilbans, and which resulted in this eighth scientific book of mine, was not supported in any way by thermoelectric organizations after the death of Stilbans. And although the technology of some thermoelectric firms like Dexter has already been brought to the nano-level, traditional thermoelectric concepts are still dominant for them. And the main interest in Local thermo-EMF is shown by electronics, which is reflected in the very title of my new book "Thermo-Photo-Electronics".

2.2 Reasons for the stagnation of the Physics of Contact Phenomena

Physics jumped from the macroscopic description

of nature to the microscopic one, skipping a whole class of NANO-phenomena [20]. And the point is not only that a very roughly constructed Quantum Theory [2,3] was used in the description of microscopic phenomena, but that at the same time, both in electronics and in optics, the model of the Ideal Infinitely Thin Interface between two media was transferred. Thus, in fact, not the contact of two media was described, but the consequences of its existence in conjugate media. Initially, the length of contact was taken into account in Optics. Both the very emergence of modern Electrophysics of Contact Phenomena, and the certainty in the very formulation of contact experiments arose only when it was understood that it had already been created, intuitively, the technology of forming an extended (of a certain size) contact with well-repeatable properties - this is a *p-n* junction [9-16].

The historical consideration of electronics without taking into account heat flows has imposed a number of restrictions on the design of devices and devices based on it. In addition, these restrictions led to a number of "theoretical" prohibitions on the existence and possibility of registration, which was revealed in the study of the Local thermo-EMF described in previous works. However, an important clarification needs to be made. Lack of understanding of the physics of Local thermodynamic effects not only imposed a ban on the Local Effects themselves, but also led to a attribution (even in WIKIPEDIA) of a primitive device to a photo-thermoelectric converter, while it contains a simple photo-conversion of the thermal radiation flux.

Similarly, under the sign "Thermo-transistor", changes in the properties of a conventional transistor from an average temperature were used, i.e. actually Local effects did not even try to use.

The main, phenomenological reason for the stagnation of the Theory of Contact Phenomena, in particular, the Theory of *p-n* junction and barrier phenomena, is, of course, that the Temperature Force has been completely thrown out of consideration. So, even in adjacent sections of the same fundamental book, thermoelectric and contact effects were considered as unrelated phenomena [5,21,22]. And this led to the fact that in thermoelectricity, Peltier heat, in violation of the Onsager Principle, continued to be attributed to the contact characteristic, and the contacts were actually considered to be infinitely thin. But precision measurements of the temperature distribution near the contact gave on the "infinitely thin" interface the inequality of asymptotics (**Figure 2**).

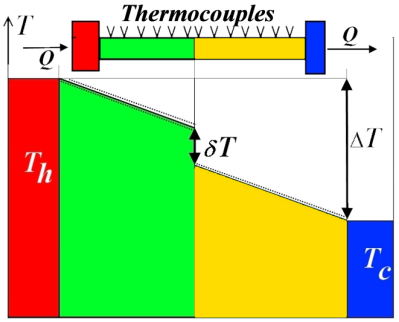


Figure 2. Asymptotic discontinuity of the temperature distribution at the interface between materials with different Fermi electron energies.

The experimentally observed temperature jump at the boundary of dissimilar media naturally follows from the production of the local Prigogine entropy at the potential barrier when current flows through it: From the absorption of energy on one side of the barrier and its release after the electron passes through the barrier. In this case, the application of thermal conductivity to the barrier is not formally applicable—the energy recovery by phonons occurs only partially (no more than half) and with a relative delay determined by the ratio of the electron transit time over the barrier to the phonon transit time through the barrier. So, the "anomalous" size dependences of thermopower and resistance shown in **Figure 1** are not anomalous, but simply not taken into account by the Theory of Contact Phenomena.

In addition, in Electronics, the Langmuir and Richardson models [23,24] are still used to calculate currents through potential barriers. These models, developed at the dawn of the last century, both on a spatial scale corresponding to the millimeter gap between the electrodes the scale of currents corresponding to the current of electrons from graphite

heated to 3000 degrees do not correspond in any way to the currents used in electronic devices. In addition, the formulas obtained within these models are very non-rigorous—the Richardson and Deshman formulas are only the first linear approximation for two special cases, giving the saturation current and a linear approximation in the region of reduced voltages less than the average thermal energy [25]. So even in serious monographs on thermionic emission these formulas are used only for decoration. And half a thousand pages are devoted to describing local patterns that have nothing to do with these formulas [26].

3. Correction & extension of electronics

Let's start the corrections with the elementary formulas for the CVC, which were obtained by Langmuir-Richardson-Deshman for Thermal Emission, i.e. not at all for semiconductor electronics, but, and they continue to be used in it:

$$J_L = \frac{\sqrt{2}}{9\pi}\sqrt{\frac{e}{m}}\cdot\frac{1}{d^2}\cdot U^{3/2},\ J_R = (1-r)AT_h^{1/2}\exp\left(-\frac{\varphi_h}{kT_h}\right) \quad (1)$$

where in the Langmuir formula, d—the barrier thickness, a U—the potential difference across it and where A—the Richardson constant, and φ_h—the work function and T—temperature (of the cathode, in thermal emission), respectively, a is the average value of the electron reflection coefficient at the boundary, which is small and, at In the analysis, we will further assume that the first parenthesis is equal to 1.

The Langmuir formula was obtained under the assumption of an initial zero velocity of all electrons above the barrier and its increment due to the electric field. And the Richardson formula—when taking into account (in the Brillouin zone) only those electrons whose velocity vector is directed towards the interface In this case, in fact, the Richardson model took into account only the difference between the electron densities at the emitter and collector above the maximum of the potential barrier (i.e., part of the concentration force) that occurs when the field is applied, multiplied by the average thermal electron velocity.

Taking into account the polarity of the electron dispersion law for a displaced barrier (**Figure 3**) allows a rigorous calculation of the ballistic current over potential barriers, the thickness of which is less than the mean free path of an electron, for electrons in the entire Brillouin zone. In this case, a strict expression for the current over the barrier can be written as a standard expression for the Richardson current with a correction factor:

$$J_R^*\left(eU^*\right) = J_R\left(eU^*\right)\cdot \mathrm{K}^{3D}\left(eU^*\right) \quad (2)$$

where the correction factor is given by expression (3) and essentially depends on the dimensionless stress reduced to thermal energy: $eU^* = \frac{eU}{kT}$.

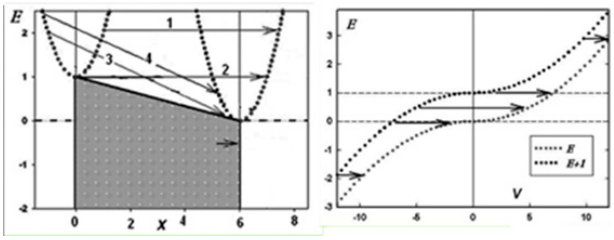


Figure 3. The spatial transition of electrons taking into account the polarity of kinetic energy (on the left—the standard dispersion law, on the right—taking into account the polarity of the current).

Taking into account additional corrections, the expression for the total current (2) turns out to be more complex than given in the work [27], but it can also be divided into factors with the coefficient of current coupling with the reduced voltage (3).

$$\mathrm{K}^{3D}\left(eU^*\right) = \frac{1}{2\left(-1+e^{eU^*}\right)}e^{-\frac{eU^*}{2}}\left(1+e^{eU^*}\right)\sqrt{\pi}\cdot \left(\begin{array}{l} 2e^{\frac{eU^*}{2}}\,HypergeometricU\left[-\frac{1}{2},0,eU^*\right]- \\ e^{\frac{eU^*}{2}}\,MeijerG\left[\left\{\{\},\left\{\frac{3}{2}\right\}\right\},\left\{\{0,1\},\{\}\right\},eU^*\right]+ \\ \sqrt{\pi}\left(BesselI\left[0,\frac{eU^*}{2}\right]+BesselI\left[1,\frac{eU^*}{2}\right]\right)eU^* \end{array}\right) \quad (3)$$

The resulting total ballistic current (2), even without taking into account the temperature force (if the temperatures of the emitter and collector are equal), shows that the Richardson formula describes, purely qualitatively, only that part of it that determines its

output to the saturation current at reduced voltages of the order of the average thermal energy (**Figure 4a**). The linear decay, borrowed in all ABC books on the Physics of Semiconductor Devices from the Richardson formula, with a decrease in the reduced voltage below the average thermal energy of electrons, as can be seen from **Figure 4**, is not observed almost to zero. This giant correction at very low voltages (**Figure 4b**) is the fundamental difference between the ballistic current and Ohm's diffuse law, which removes the Langmuir-Richardson prohibition on their measurability of currents at low voltages. A significant, not only quantitative, but also qualitative correction for the ballistic current in terms of saturation current and at voltages is greater than the average thermal energy.

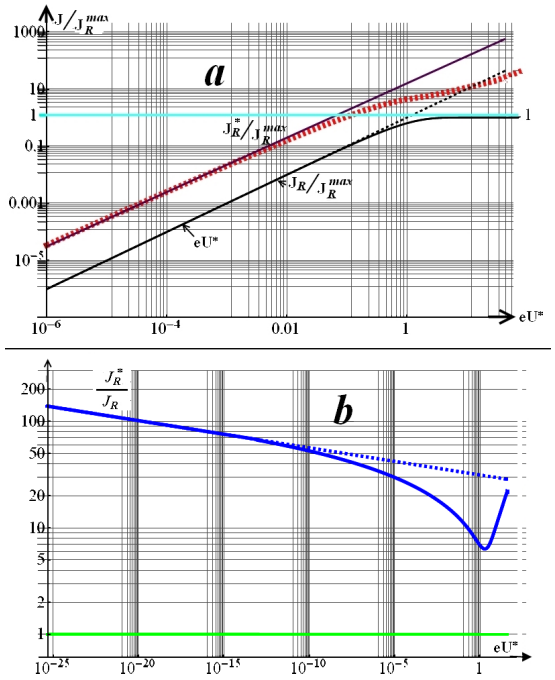


Figure 4. Dependence of the Richardson current and the total ballistic current (Richardson) on the reduced voltage: a—currents in relation to the Richardson saturation current, b—ballistic correction factor to the Richardson current.

Of course, there are limitingly measurable currents, but they are by no means limited to reduced voltages of the order of the average thermal energy of electrons, which characterizes only Ohm's diffuse law. The minimum measurable currents are determined by the signal-to-noise ratio, which depends on the registration time: $kT = U \cdot J \cdot \Delta t_{\text{measurement}}$, and in the limit—by the Heisenberg Uncertainty Principle. But, in fact, the neglect in the p-n junction of currents less than those corresponding to this voltage, which is a rough model limited the linear section of the current-voltage characteristic (theoretically, they were measured experimentally for a long time) and became one of the components of the PROHIBITION of the existence of Local thermo-EMF.

A significant, not only quantitative, but also qualitative correction of the ballistic current was obtained for the saturation current, for current and at voltages greater than the average thermal energy. As can be seen from **Figure 4b**, at reduced voltages greater than the thermal energy of electrons, an avalanche-like increase in current occurs—pre-breakdown.

Honest experimenters carry out precision measurements and trust their results, rather than theoretical prohibitions, which are canonized, but sometimes reflect only some local patterns. But in Science, this happens, unfortunately, often—the Ideas of its true Creators are not fully understood, but are picked up by opportunists (market players)—developers. So these Ideas are actively promoted by developers, but with distortions and errors. So it was, for example, with many operating devices created by Geniuses: Leonardo da Vinci, Nikola Tesla and Lev Theremin, who could not reproduce later. And now, when the fundamental sections of Physics have reached the modern level, it has become clear that an entire industry has been formed—Electronics, the instrumental and technological problems of which are related to the fact that the theory of the p-n transition is built in violation of the laws of non-equilibrium thermodynamics and that this is largely determined by the distortion of Ideas, which the founders of electronics, Losev and Tauc, came to purely intuitively (the Nobel laureates for the transistor themselves called the creator of Losev's resistance transformer his father).

So, it was with the p-n junction, first created in silicon carbide and described in the 30s of the last century by Oleg Losev. Losev himself immediately intuitively realized that this was a current device. And he was able to use it almost immediately: He

created on it an LED, a photodiode, and a resistance transformer, which the Nobel laureates called in short (in English) a transistor and even a transistor receiver. But the physics of the p-n junction and the description of the operation of the listed devices based on the p-n junction were built by analogy with a radio tube, which, in principle, is a field device. Thus, when solving various problems for semiconductor devices in the p-n junction, the cause-current and the effect-voltage were rearranged. And the prominent physicist Abram Fedorovich Ioffe did not fully understand Losev then, who was half a century ahead of modern electronics. But academician Ioffe, not like the current "luminaries" of science, achieved the assignment of a candidate of physics and mathematics to him. And only after almost 100 years, the return to Losev's current circuit made it possible to significantly improve the characteristics of semiconductor devices.

Tauc is another, Czech Corypheus, ahead of his time. He was the first, immediately after the liberation of Prague by the Soviet troops, to establish the production of point transistors in Prague (formerly of the Bell Company). And he immediately discovered thermoelectric effects in the p-n junction and honestly described them. But then (and even now) thermoelectricity itself was stuck at the macroscopic, purely diffuse level of describing the phenomenon and classified Tauc's results as anomalies.

The description of the generation of photo-EMF in the *p-n* junction was also carried out for a field device. Photo-EMF, in principle, was correctly associated with a potential barrier and an electric field in it, as a force, but balancing only the concentration force. In this rough approximation, the p-n junction diagram shows the complete alignment of the Fermi levels of the electron and hole regions (**Figure 5**) with the formation of a region with full compensation of electrons and holes (i-region).

The thermodynamic discovery by Ilya Prigogine of the production of local entropy helped to restore the correct description of the physics of the p-n junction. It made it possible to understand that in the p-n junction, described for the reasons noted above, within the framework of a truncated concentration-electric phenomenology, it is necessary to use an extended phenomenology supplemented by a heat flux [11-16].

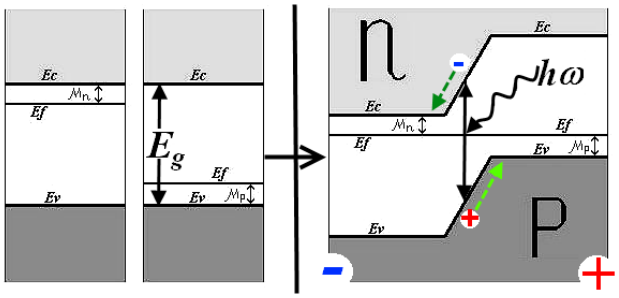


Figure 5. The traditional scheme for the transformation of energy bands of semiconductors with different types of conductivity upon their contact and the opening of the p-n junction formed in this case when the i-region is irradiated with light.

From the balance of thermodynamic flows ***J***, the equality of the concentration force FN and the electric force ***FE*** in the p-n junction naturally follows. But we will not dwell on this refinement of the description of the photoelectric effect for the time being. The analysis of thermoelectric and thermionic effects showed that from the general phenomenology, strictly corresponding to the Curie theorem and the principle of symmetry of the Casimir-Onsager kinetic coefficients, in the description of the *p-n* transition and, thereby, the photoelectric effect, the temperature force ***FT*** is excluded from consideration of thermodynamic flows ***J*** (**Table 1**).

Table 1. Three private phenomenologies, traditionally using only two thermodynamic forces in three different branches of Physics.

E-T phenomenology $J_E=\sigma E+\sigma(S\Delta T)$ $J_T=\prod J_E+K\Delta T$	Thermoelectricity $J_E=L_{EE}F_E+L_{EN}F_N$ $J_T=L_{ET}F_E+L_{TT}F_T$
E-N phenomenology p-n junction	Volt-Amp Characteristic (VAC) $J_E=L_{EE}F_E+L_{EN}F_N$ $J_N=L_{EN}F_E+L_{NN}F_N$
T-N phenomenology Vacuum gap	Thermal emission $J_T=L_{TT}F_T+L_{NT}F_E$ $J_N=L_{TN}F_E+L_{NN}F_N$

Table 1 shows the Phenomenologies underlying the incomplete description of a number of effects as Fragments of the General Phenomenology (for thermoelectricity, the system of equations, in addition to

the canonical form, is also written using traditional coefficients: electrical conductivity σ, thermal conductivity **K**, Seebeck **S** and Peltier **Π**).

At the same time, the standard band structure of the junction itself is modified, taking into account the temperature force, and the energy diagram of the equilibrium p-n junction is modified, which, without taking into account the temperature force, gave, as shown in **Figure 5**, the potential barrier value approximately equal to the band gap of the semiconductor. Taking into account the temperature force, in the absence of heat flux, the value of the potential barrier turns out to be equal to half the band gap (**Figures 6a-6b**). And when it turns on the heat flow, the value of the potential barrier will increase until a tunnel breakdown occurs (**Figure 6c**).

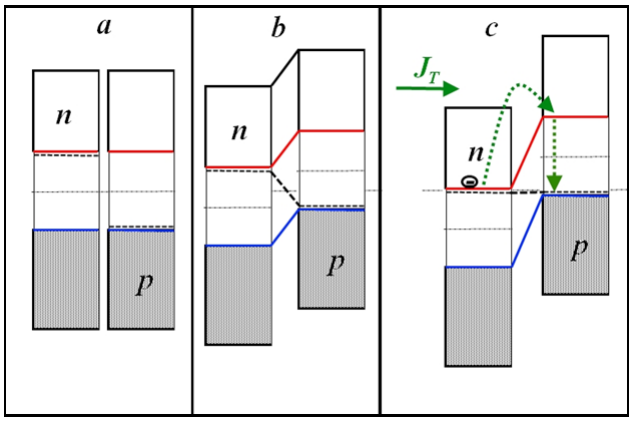


Figure 6. Energy diagrams of two semiconductors of different types of conductivity: a—before bringing them into contact, b—equilibrium state after their contact at equal temperatures, c—equilibrium state with heat flow through the p-n junction.

As shown in **Figures 6a and 6b**, between the semiconductors brought into contact in an equilibrium state, due to the balance of electric and concentration forces, the difference in electric potentials at the contact (the red line is the bottom of the conduction band, the blue line is the top of the valence band) is equal in magnitude and opposite in sign to the difference at the boundary of the concentration potential (dashed black line). Thus, given in **Figure 5**, in accordance with the complete system of equations for thermodynamic forces and flows, it already allows eliminating the theoretical equal 2 in the description of the transition without a temperature gradient, which was associated with an empirical coefficient due to the imperfection of materials. The potential difference across the transition plates is equal in this case to half the band gap. And the equality of concentration potentials corresponds to the Local Thermo-EMF and occurs, of course, only with a heat flow through the transition (**Figure 6c**).

In this case, the current-voltage characteristic (CVC) of the p-n junction, of course, depends on the temperature difference on its plates, which is determined by the heat flux through the p-n junction. If, in accordance with Losev's theory, we take into account the primacy of the current in the p-n junction, then the contribution of the temperature force gives a shift in the CVC, described by the formula:

$$\Delta J_{Th/Ph} = k_{Th//Ph} \cdot O_{Th/Ph} \quad (4)$$

where $O_{Th/Ph}$ the energy flow through the p-n junction, and $k_{Th//Ph}$ the current coupling coefficient in the *p-n* junction from the heat or light flux (for light—the quantum yield at a given wavelength of light).

There is a current-voltage characteristic shift, similar to the photo-effect shift, but of the opposite sign EMF (**Figure 7**).

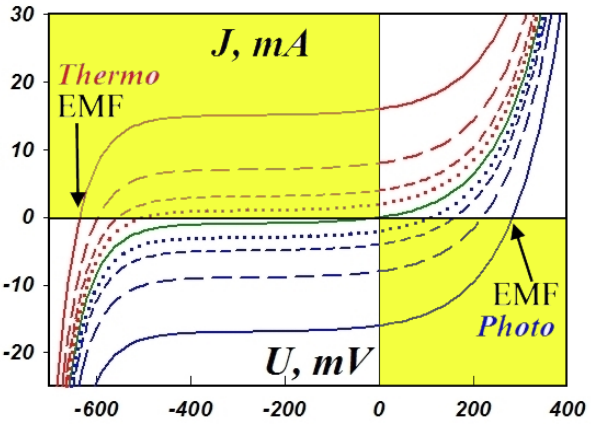


Figure 7. Calculated shifts in the current-voltage characteristic of an ideal p-n junction by equivalent (in terms of power) flows: thermal (red curves) light (blue curves), with an increase in flows by 2 times. Generation quadrants are marked in yellow.

The experiments confirmed the qualitative picture of the displacement of the CVC by a constant, stabilized heat flux (**Figure 8**).

But as dynamic experiments have shown, when

a stabilized but modulated heat flux is applied, the shift of the transition by a constant stabilized current leads to a change in the frequency response with its passage, in accordance with "Experimental and Theoretical Expansion of the Phenomenology of Thermoelectricity" [12], through the Gaussian thermoelectric resonance, which is determined by the phase reversal of the thermal signal (**Figures 9a, 9b, 9c**).

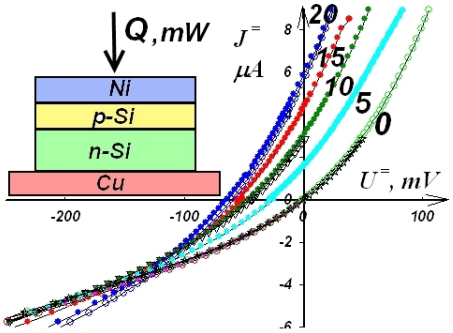


Figure 8. Experimental current-voltage characteristics of a p-n junction with stabilized different heat fluxes through it.

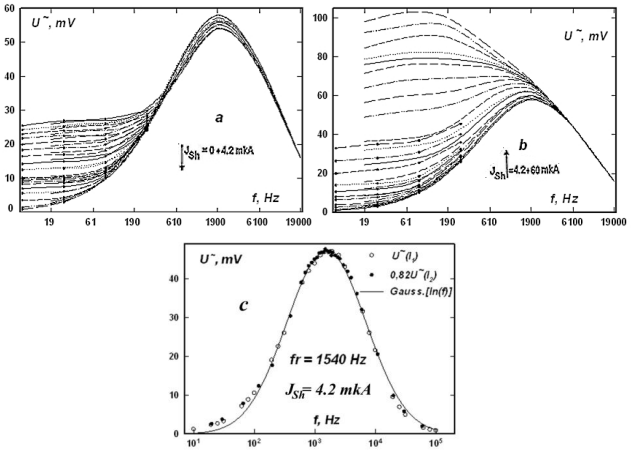


Figure 9. Dependence of the frequency response of the p-n junction on its DC bias [28]: a—decrease in the low-frequency branch with an increase in the bias current to a critical value of 4.2 μA, b—increase in the low-frequency branch with an increase in the bias current above the critical value, c—Thermo-Photo-Electric resonance at a critical value of the bias current, described normally by the Gauss dependence from the logarithm of the frequency.

So the differences in the I-V characteristics in **Figures 7 and 8**, in principle, not only do not violate the overall picture, but they detail it as the presence of two antiphase signals determined by technological and design factors.

As shown in **Figures 7 and 8**, when a heat flow flows through the p-n junction, an area of positive currents arises in its I-V characteristic in the upper left quadrant at negative voltages at the p-n junction, which, in full accordance with the concepts of generators (the simplest is an electric battery), is the area of generation of electrical energy due to the flow of heat flowing through the transition (**Figure 10**).

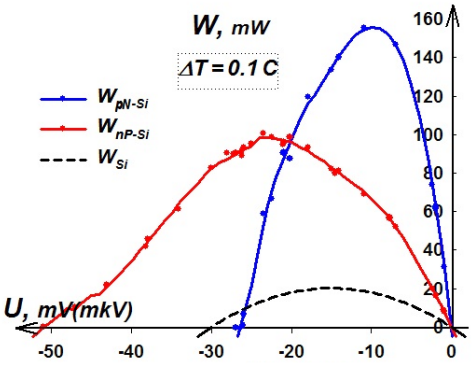


Figure 10. Experimental thermo-generator characteristics of p-n junctions and optimally doped silicon (for silicon, the optimal EMF—μV is three orders of magnitude lower than that of p-n junctions—mV).

This was experimentally confirmed on the developed detectors based on Local Thermo-EMF in silicon junctions, the volt-watt sensitivity of which was obtained three orders of magnitude higher than that of detectors based on the traditional macroscopic Seebeck effect.

At the same time, the extended phenomenology shows that there can be complex, three-phase effects in the *p-n* junction, which makes it possible to optimize, in particular, the combination of local thermoelectric effects with photoelectric effects in it.

Previously, the generator characteristic of the transition was constructed without taking full account of the specifics of the equilibrium state, taking into account the local production of Prigogine's entropy. In this case, zero current through the junction corresponds to a zero temperature drop. But in fact, the p-n junction is Maxwell's demon—a "gear" wheel separating hot and cold current carriers, in principle, allowed at the micro level by the production of Prigogine's Local Entropy. And the mechanism of its operation is obvious from **Figure 5**: When semiconductors come into contact, its asym-

metry (polarity) occurs. At the same time, to start the electron transfer process, an energy equal to half the band gap is sufficient, while after their transfer and annihilation with holes, their reverse transfer requires an energy equal to the full band gap. So, a local temperature drop occurs on the plates of the p-n junction, which determines the local thermodynamic equilibrium. In this case (**Figure 10**), the experiment shows that the CVC has a shape fundamentally similar to that shown in **Figure 8**.

The sign of the Local thermo-EMF arising in the p-n junction does not depend on the polarity of the heat flux, as in the case of voltages determined by the diffuse Seebeck coefficient (which are 3 orders of magnitude less), but is completely determined, in accordance with the displacement of the CVC (**Figure 7**) by the heat flux modulus of polarity p-n transition, i.e. diffusion of the main carriers of its hot lining (**Figure 11**).

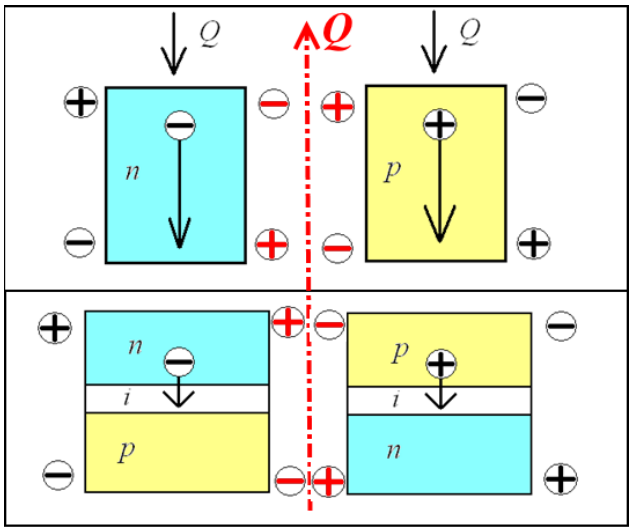


Figure 11. Dependence of the polarity of the EMF on the direction of the polarity of the heat flux for conventional diffuse thermoelectrics (top) and the dependence of the polarity of the Local EMF only on the polarity in the p-n junction (bottom).

On a structure of series-connected *p-n* junctions with a thickness of 3 μm, grown on a 300-micron device silicon substrate, a similar independence of the generated temperature drop from the current polarity is observed at short recording times. And at long measurement times, only a certain asymmetry is observed, which is associated with the shunting effect of a thick substrate (**Figure 12**).

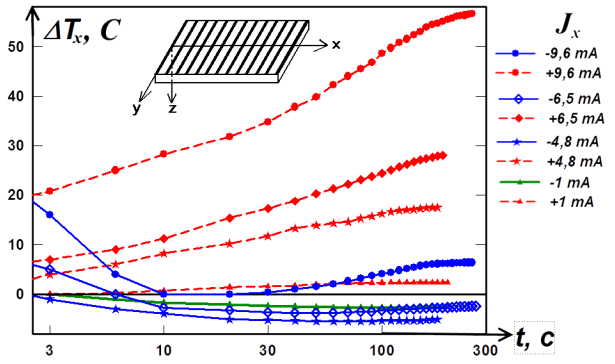


Figure 12. Separation of an analogue of Peltier heat by a structure of series-connected p-n junctions grown on a silicon substrate [28-31].

As can be seen from **Figure 12**, the fast temperature response, determined only by the polarity of a three-micron structure of series-connected *p-n* junctions, has a slow temperature effect, determined by a thick silicon substrate.

It is also characteristic that the thermal shunting of the structure of p-n junctions completely linearizes the CVC, which directly indicates that the heat released by the structure is determined by the Local thermo-EMF.

The use of extended concentration-electrical-thermal phenomenology made it possible both to confirm the current (primacy of the current) of the Losev *p-n* junction, and to transfer the "anomalous" thermoelectric power discovered by Tauc in the p-n junction to the category of normal-local, and to describe new experimental results of studies of contact thermoelectric power. In addition, the extended phenomenology made it possible to understand that macroscopic thermoelectricity is artificially limited only by diffuse thermoelectric materials and showed that for diffuse thermoelectrics, the efficiency of thermoelectric conversion achieved in practice is already close to the theoretical limit.

In addition, extended phenomenology has shown that the efficiency of thermoelectric conversion based on local thermoelectric power does not have a diffuse limit and can be dramatically increased by several times compared to that achieved using the Seebeck effect in traditional diffuse materials (**Figure 10**).

This was experimentally confirmed on the devel-

oped detectors based on Local Thermo-EMF in silicon junctions, the volt-watt sensitivity of which was obtained three orders of magnitude higher than that of detectors based on the traditional macroscopic Seebeck effect.

At the same time, the extended phenomenology shows that there can be complex, three-phase effects in the p-n junction, which makes it possible to optimize, in particular, the combination of local thermoelectric effects with photoelectric effects in it.

4. Reversible effects

Since what differs from the generally accepted one often "does not fit in the head" and is perceived as erroneous or mystical, an ELEMENTARY explanation is required for the material shown in **Figure 11** and **Figure 12**.

Reversible Processes, forbidden by the Second Law of Thermodynamics, connected with the transformation of Chaos into Harmony (the impossibility of this transformation in inanimate Nature), we will not consider here [32-35].

"The law of non-decreasing entropy, or the so-called physical meaning of the second law of thermodynamics, was discovered by Rudolf Clausius (1865), and its theoretical justification was given by Ludwig Boltzmann (1870s)."

But we will consider not just Reverse Processes, such as a change in the sign of a mechanical or electrical force with a corresponding change in the sign of movement, as in the same linear Ohm's Law, but we will consider exactly the Reverse Processes, which, in principle, are not prohibited by the Second Law of Thermodynamics, but simply limited in effectiveness. The simplest examples are the conversion of mechanical or electrical energy into heat, which the Heat Engine has long "learned" to convert back into mechanical or electrical energy. And the limit of effectiveness of this Reverse Transformation, as follows from the Second Law of Thermodynamics itself, is determined by the Carnot cycle. One of these Thermal Engines is a thermoelectric generator, the efficiency of which, in principle, is determined by the same Carnot cycle, but in reality it is less than three times, since the diffuse Seebeck effect is used, i.e. friction is used squared. In this regard, Local thermo-EMF, using not diffuse, but ballistic effects, eliminates the friction multiplication, thereby making it possible to approach the Carnot cycle. But thermoelectricity itself also demonstrates the Reverse Process—the flow of electrical energy (into an external circuit) from a thermogenerator can, in principle, be opposed to the flow of heat from a thermoelectric refrigerator.

And in this regard, there is nothing supernatural in the fact that with the help of Local thermo-EMF it is possible to generate electricity both due to the flow of the body, and to generate a heat flow due to electric energy.

In considering Local thermo-emfs, we have already started from the band diagram and from the current-voltage characteristics, similar to those used to describe the photo-effect. And for the photo effect, the reverse effect used in LEDs is also well known. In principle, these both effects are also the implementation of the Heat Engine, simply because of the high light temperature, which acts as a hot temperature in the Carnot formula, it is believed that their Carnot coefficient is close to unity and, at the same time, all energy processes in the p-n junction can be fully described by the change in the electron energy in the transition diagram. In practice, this is far from the case, and in terms of efficiency, photocells only approach the efficiency of an internal combustion engine. And the LEDs heat up so much that they even burn out with improper cooling. And when indirect-gap semiconductors are used as a material for these diodes, the efficiency also decreases further due to the fact that the transition of an electron from one energy level to another is carried out not only due to a quantum of light, but also due to thermal vibrations of phonon atoms.

Nevertheless, their principles of operation in an ideal diode can be described in the "language" of light-electron.

For a photodiode, as shown in **Figure 7**, the light flux shifts the current-voltage characteristics of an ideal diode. At the same time, in the photo-quadrant

marked in yellow, the current and voltage are in antiphase and the corresponding power is negative (**Figure 13**).

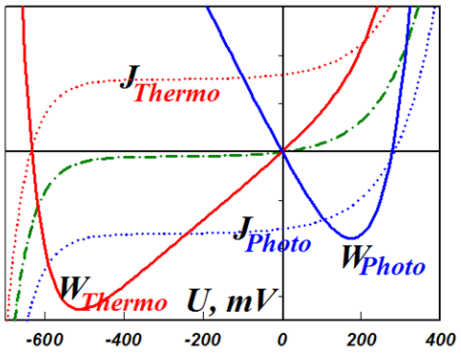


Figure 13. Anti-phase current and voltages give negative (generator) power both with the photo effect (blue color) and with the Local thermal effect.

As shown in **Figure 13**, we obtain a similar negative power peak for the thermal quadrant, where the current and voltage are in antiphase (in experimental **Figure 10**, the corresponding thermal power peaks are shown positive for convenience). It is like a battery delivering positive power to an external circuit, the current and voltage inside it are also in antiphase, which corresponds to the removal of this power from the battery to transfer it to the external circuit.

The light effect, the reverse of the photo effect, has been studied quite well. Therefore, we will begin the analysis of Reverse Effects with him. As was customary in all Electronics, and analyze the light effect as a field effect, determined by the achievement of a certain critical voltage on the p-n junction on the direct branch of the CVC. But if we take into account the linear relationship of different flows (formula 4), then the experimental CVC of the direct branch of the LED corresponds to positive current displacements of the direct branch of its CVC, which corresponds to the expansion of the phenomenology of the description of the p-n junction when taking into account the light flux (**Figure 14**, on the right).

So, taking into account the primacy of the current, it is possible to construct a differential I-V characteristic of an ideal diode, displaced, as in the case of the photo effect, by the opening (of a different polarity) current, determined by the emitted light flux. So for an ideal diode, we will get the power released in the p-n junction near the critical voltage, only in the form of light. For a real diode, for the reasons noted above, only a part of the power will be released in the form of light, and, at high currents, the light will be extinguished. All the electrical power consumed by the diode will be released as heat. Such a current analysis of the light effect makes it possible to clarify the description of the direct photo effect as well. As shown in **Figure 5**, the photo effect, due to the smallness of the photon momentum, is associated precisely with vertical transitions of electrons. But the critical voltage of the light of the diode is reached when the spatial transition of electrons becomes possible (**Figure 3**), which, at the same time, strictly corresponds to the law of conservation of momentum and does not require the participation of phonons in direct-gap semiconductors.

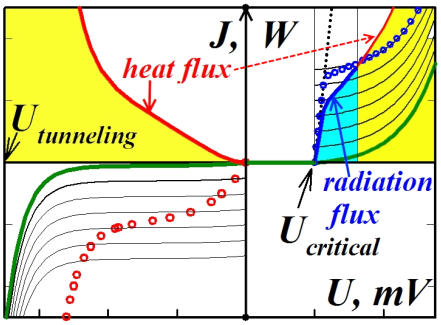


Figure 14. Schematic diagram of the description of the forward and reverse branches of the I-V characteristics (shown by circles, blue for the light effect, red for the Local thermal effect in the p-n junction) by the current bias of the I-V characteristics of an ideal diode.

When the ideal diode is displaced by the current in the shutoff direction, the electrical energy consumed by the diode in the initial section of the I–V characteristic (**Figure 12**) will in fact similarly turn into a heat flux, which, as in the light effect, corresponds to the displacement of the I–V characteristic of the ideal diode, but in terms of the shutoff current (**Figure 14**, left).

And in both considered cases, in the initial sections of the CVC, the parameters of the p-n junction plates will determine on which plate the release of light or heat will occur more. And with an increase in currents above certain critical ones, heat will be released purely ohmic-isotropically. And this is the

boundary between flows and isotropic selection.

Strict, quantitative consideration requires, as shown above (**Figure 6**), taking into account the "Maxwell's demon" in the initial CVC, taking into account that the equilibrium state of the p-n junction corresponds to a small temperature difference between its plates. But this temperature drop is associated with specific details of the p-n junction and will not be considered here. We only note the main thing - a strict phenological description, developed, in principle, for any number of independent forces (***J***) and flows (***F***), as shown by studies on the NANO-scale, requires taking into account not three (as shown in **Table 1**), but four pairs of forces and flows: electrical, concentration, thermal and light.

$$\begin{aligned} J_E &= L_{EE}F_E + L_{ET}F_T + L_{EN}F_N + L_{EPh}F_{Ph} \\ J_N &= L_{NE}F_E + L_{NT}F_T + L_{NN}F_N + L_{NPh}F_{Ph} \\ J_T &= L_{TE}F_E + L_{TT}F_T + L_{TN}F_N + L_{TPh}F_{Ph} \\ J_{Ph} &= L_{PhE}F_E + L_{PhT}F_T + L_{PhN}F_N + L_{PhPh}F_{Ph} \end{aligned} \quad (5)$$

where L is the corresponding 16 kinetic coefficients related to each other, as follows from the "reversibility"—the reversibility of effects (**Figures 13-14**), by the Onsager symmetry principle, which reduces the number of independent kinetic coefficients to 10:

$$\begin{aligned} J_E &= L_{EE}F_E + L_{ET}F_T + L_{EN}F_N + L_{EPh}F_{Ph} \\ J_N &= L_{EN}F_E + L_{NN}F_N + L_{NT}F_T + L_{NPh}F_{Ph} \\ J_T &= L_{ET}F_E + L_{NT}F_N + L_{TT}F_T + L_{TPh}F_{Ph} \\ J_{Ph} &= L_{EPh}F_E + L_{TPh}F_T + L_{NPh}F_N + L_{PhPh}F_{Ph} \end{aligned} \quad (6)$$

By solving this system for any flow (force) and using the boundary conditions, one can obtain the maximum efficiency of electric power and heat flux generation [11-13], or the maximum efficiency of light generation. Moreover, it is possible to achieve the maximum efficiency of the transistor and thereby reduce its own noise level and increase its speed.

Only in this case it must be borne in mind that these are Local Kinetic coefficients that characterize a non-homogeneous material, an artificially created "Maxwell's demon"—an element of a working nano-structure, in particular, the ideal diode described by us.

In semiconductor devices with weak light and photo effects, to find the maximum efficiency, one can limit oneself to three simplified systems of equation (6) and analyze the three-dimensional thermo-voltage-ampere surface (**Figure 15**).

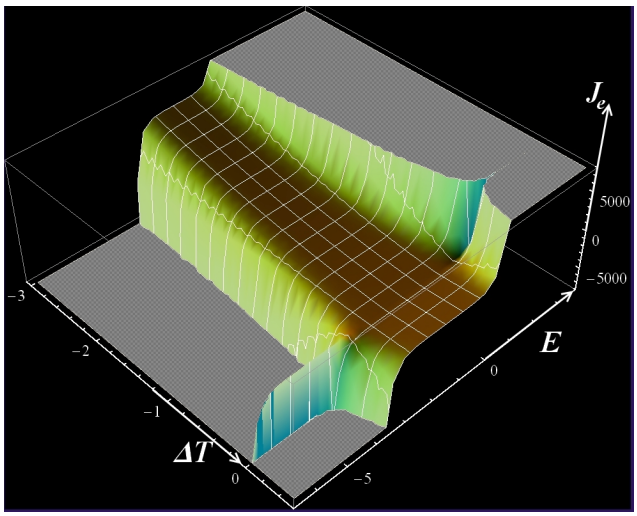


Figure 15. CVC expansion takes into account the temperature force.

In the general case, it is necessary to analyze a 4-dimensional thermo-photo-voltage-ampere surface.

5. Conclusions

The registration of the Effects, which were classified as quantum, as if by itself, implied their measurability. Additionally, this was supported by hastily made Quantum Statistics, which, as it were, fenced off the measurability of Quantum Effects from its limitations by Classical Thermodynamics (how loosely this was done will be shown in the chapter of my future work "Quantum Extension of Classical Representations" [36]). But, even without the macroscopic parameters calculated from the mystical Schrödinger wave functions, simply from the analysis of Newton's Elementary Particle, it is possible to obtain both their measurable diffuse light quanta flux, and the measurable wave of coherent laser light.

Local—missed between macroscopic and microscopic NANO-effects, in this regard, were less fortunate, because they directly adjoined, and in thermoelectricity they were determined by Thermodynamics and its limitations simply forbade their measurability. And these restrictions-prohibitions were lifted by

Ilya Prigogine's own Thermo-Dynamics of flows, which transferred the science itself, previously called Thermodynamics, to the category of Thermo-Statics.

At the same time, the study of Local Effects raised a number of fundamental questions that turned out to be in the "fork": From—this cannot be, to the same—elementary. This also applies to the registration of the Local thermo-EMFs discovered and previously classified as anomalous, but not only—the entire Physics of semiconductors contains a systematic elementary error in Phenomenology itself.

But by and large, not a primitive, but an elementary solution allows you to create devices with fundamentally (drastically) improved characteristics. This was the case before, when new devices were created based on the well-known First Principles, the same was confirmed when creating detectors based on the discovered NANO-effects, the study of which resulted in the previously missed NANO-Physics.

Thus, in order to achieve this fundamental improvement in the characteristics of devices and devices, it was necessary to eliminate this fundamental error and solve a number of fundamental issues, a number of methodological measurement issues, and a number of issues related to the technology of manufacturing efficient semiconductor nanostructures, and a number of design issues.

Within the framework of this study, answers were obtained to a number of fundamental questions.

1) Measurability of Local Effects.

2) The primacy in the effects of the current as the cause of the effect and the secondary nature of the voltage, as a consequence, taking into account external conditions.

3) Electric currents over potential barriers in semiconductors.

4) Thermal force in the p-n junction.

5) Extended Thermo-Photo-phenomenology for describing effects in a Solid State.

Conflict of Interest

There is no conflict of interest.

References

[1] Ordin S.V., 2018. C & BN-foundation for atomic-crystalline orbitals. Global Journal of Science Frontier Research—Physics & Space Science. 18(5), 17-47.

[2] Ordin, S., 2022. Gaps and errors of the Schrödinger Equation. Global Journal of Science Frontier Research. 22(3), 1-5.

[3] Ordin, S.V., 2021. Foundations of Planck-Einstein quantization (Thematic collection of recent studies reviewed in scientific journals). LAP Lambert Academic Publishing: London.

[4] Ordin, S., 2022. Foundations of quantization principles. Jenny Stanford Publishing Pte Ltd.: Singapore.

[5] Pickus, G.E., 1957. Контактные явления. Полупроводники в науке и технике, т.1 (Russian) [Contacts phenomena. Semiconductors in a science and engineering, v. I]. Academy of Sciences USSR: Moscow. pp. 113-132.

[6] Jan Tauc, 1962. Photo—and thermoelectric phenomena in semiconductors. (translation from Czech). Foreign Literature: Moscow.

[7] Ordin, S.V., 1997. Optimization of operating conditions of thermocouples allowing for nonlinearity of temperature distribution. Semiconductors. 31(10), 1091-1093.

[8] Onsager, L., 1931. Reciprocal relations in irreversible processes. I. Physical Review. 37(4), 405.

[9] Ordin, S.V., Zjuzin, A.J., Ivanov, Y. (editors), et al., 2010. Nano-structured materials for thermoelectric devices. The 29th International Conference on Thermoelectrics; 2010 May 30-Jun 3; Shanghai.

[10] Ordin, S.V., Wang, W.N., 2011. Thermoelectric effects on micro and nano level. Journal of Advances in Energy Research. 9, 311-342.

[11] Ordin, S.V., 2017. Refinement and supplement of phenomenology of thermoelectricity. American Journal of Modern Physics. 6(5), 96-107.

[12] Ordin, S.V., 2018. Experimental and theoretical expansion of the phenomenology of thermoelectricity. Global Journal of Science Frontier

Research-Physics & Space Science (GJSFR-A). 18(1), 1-8.

[13] Ordin, S.V., 2017. Cardinal increase in the efficiency of energy conversion based on local thermoelectric effects. International Journal of Advanced Research in Physical Science. 4(12), 5-9.

[14] Ordin, S.V., 2018. Anomalies in thermoelectricity and reality are local thermo-EMFs. Global Journal of Science Frontier Research-Physics & Space Science (GJSFR-A). 18(2), 59-64.

[15] Ordin, S.V., Zhilyaev, Y.V., Zelenin, V.V., et al., 2017. Локальные термоэлектрические эффекты в широкозонных полупроводниках(Russian) [Local thermoelectric effects in wide-gap semiconductors]. Semiconductors. 51, 883-886. DOI: https://doi.org/10.21883/FTP.2017.07.44643.29

[16] Ordin, S., 2023. Foundations of polymer thermoelectronics. Journal of Materials and Polymer Science. 3(1), 1-5.

[17] Ordin, S., 2015. Достижения и проблемы термоэлектричества, опубликовано (Russian) [Achievements and Problems of Thermoelectricity]. Available from: http://rusnor.org/pubs/articles/12707.htm

[18] Ordin, S.V. (editor), 1997. Peltier heat as a volume property and optimization of working regimes of thermoelements in real conditions. ICT'97, XVI International Conference on Thermoelectrics; 1997 Aug 26-29; Hotel Art'otel Dresden. USE: IEEE. p. 96-97. DOI: https://doi.org/10.1109/ICT.1997

[19] Ordin, S.V., 2018. Anomalies in thermoelectricity and reality are local Thermo-EMFs. Global Journal of Science Frontier Research-Physics & Space Science (GJSFR-A). 18(2), 59-64.

[20] Ordin, S., 2012. НАНО или новый образ мышления, опубликовано (Russian) [NANO or a new way of thinking]. Nanotechnological Society of Russia. Available from: http://rusnor.org/pubs/articles/7556.htm

[21] Stilbans, L.S., 1957. Термоэлектрические явления. Полупроводники в науке и технике, т.1 (Russian) [The thermoelectric phenomena. Semiconductors in a science and engineering, v. I]. Academy of Sciences USSR: Moscow. pp. 113-132.

[22] Ziman, J., 1974. Principles of the theory of solids. Cambridge University Press: Cambridge.

[23] Sze, S.M., 1981. Physics of semiconductor devices. Wiley: New York. pp. 365.

[24] Gamache, V.I., 2000. Физика полупроводниковых устройств, Второе Издание (Russian) [Physics of semiconductor devices, Second Edition]. NTL Publishing House: Moscow. pp. 426.

[25] Ordin S.V., 2014. Баллистическая модель движения электронов над потенциальным барьером, ФТИ им. А.Ф. Иоффе, Российской Академии Наук, Санкт Петербург, Россия (Russian) [Ballistic model of the movement of electrons over potential hill, PHTI of A.F. Ioffe of the Russian Academy of Sciences, St.-Petersburg, Russia]. Interstate Conference: Thermoelectrics and Their Application. p. 199-203. [cited 2014 Nov 14]. Available from: http://rusnor.org/pubs/articles/11583.htm

[26] Moyzhes, B., Pikus, G., 1973. Термоионное преобразование под редакцией Б.Я (Russian) [Thermionic converters and low-temperature plasma]. Academic of Sciences USSR: Moscow. pp. 532.

[27] Ordin, S.V., 2021. Allowed spatial transitions and cancellation of the Richardson-Langmuir ban. Non-Metallic Material Science. 3(1), 15-23.

[28] Ordin, S.V. (editor), 2002. Contact thermopowers. Proceedings of ICT'02 XXI International Conference on Thermoelectrics; 2002 Aug 25-29; Hyatt Regency Hotel, Long Beach. USA: IEEE.

[29] Ordin, S.V., Sokolov, I.A., Zjuzin, A.J. (editors), 2006. Термоэлектрические процессы в p-n переходах (Russian) [Thermoelectric processes in p-n junctions]. Works of X Interstate Seminar: Thermoelectrics and Their Application, A.F. Ioffe PhTI of the Russian Academy of Sciences; 2006 Nov 14-15; St.-Petersburg. p. 41-47.

[30] Ordin, S.V., 2020. Local (NANO) thermoelectric effects. The Journal of Modern Technology and Engineering. 5(1), 107-109.

[31] Ordin S.V. (editor), 2015. Оптическая методика измерения локальных термо-ЭДС, ФТИ им. А.Ф.Иоффе Российской Академии Наук, Санкт Петербург, Россия (Russian) [Optical technique of measurement local thermo-EMF, PHTI of A.F.Ioffe of the Russian Academy of Sciences, St.-Petersburg, Russia]. Interstate Conference: Thermoelectrics and Their Application. p. 234-237.

[32] Ordin, S., 2017. Refinement of basic physical models. Lambert: London. pp. 82.

[33] Ordin, S., 2022. Характеризация Состояния Материи, опубликовано (Russian) [Characterization of the state of matter]. Nanotechnological Society of Russia. Available from: http://rusnor.org/network/social/user/10216/blog/3877/

[34] Ordin, S., 2022. Степени Гармонии Природы, опубликовано (Russian) [Degrees of harmony of nature]. Nanotechnological Society of Russia. Available from: http://rusnor.org/network/social/user/10216/blog/3894/

[35] Ordin, S., 2023. Foundations of thermoelectronics. International Journal of Physics and Mathematics. 5(1), 15-19.
DOI: https://doi.org/10.33545/26648636.2023.v5.i1a.46

[36] Ordin, S., 2023. Correction and extension of quantum statistics. Science Set Journal of Physics. 2(2), 1-4. Available from: https://www.mkscienceset.com/articles_file/246-_article1682576762.pdf

BILINGUAL PUBLISHING GROUP

Tel:+65 65881289
E-mail: contact@bilpublishing.com
Website:https://journals.bilpubgroup.com

2661-3204
01
9 772661 320230