The Radio Science Bulletin

ISSN 1024-4530

INTERNATIONAL
UNION OF
RADIO SCIENCE

UNION
RADIO-SCIENTIFIQUE
INTERNATIONALE

2008 Awardees

Prof. W.J. Welch

Prof. A.E.E. Rogers

Prof. U.S. Inan

Prof. H. Matsumoto

Dr. D.F. Sievenpiper

No 325
June 2008

Publié avec l'aide financière de l'ICSU
URSI, c/o Ghent University (INTEC)
St.-Pietersnieuwstraat 41, B-9000 Gent (Belgium)

Contents

Front cover: At the XXIXth URSI General Assembly in Chicago (USA) this August, the scientists whose pictures feature on the front cover were presented with the URSI Awards. For more information, please turn to page 10 of this Bulletin.

For information, please contact :
The URSI Secretariat
c/o Ghent University (INTEC)
Sint-Pietersnieuwstraat 41, B-9000 Gent, Belgium
Tel.: (32) 9-264 33 20, Fax: (32) 9-264 42 88
E-mail: info@ursi.org
http://www.ursi.org

Editorial

Special Sections

This issue and the December issue of the *Radio Science Bulletin* contain special sections honoring Jenifer Haselgrove and her work on ray tracing. This led to the Haselgrove Equations, which have found wide application in radio science. A brief summary of Haselgrove's career is included in the introduction to the special sections. Also included there are introductions to the papers that appear in this issue, and a summary of the papers that will appear in December.

The Guest Editors for the special sections are Rod Barnes and Phil Wilkinson. Their efforts, and those of the authors who have contributed such very interesting papers, are greatly appreciated.

GPS and Time

We have one additional paper in this issue, by P. Banerjee. He provides a review of the use of the Global Positioning System (GPS) for time dissemination and time transfer. He begins with a brief history of time-transfer techniques, focusing on those associated with navigation systems. He then explains how GPS disseminates time, and two basic methods of using GPS to determine and transfer time. A discussion of the various factors that affect GPS accuracy follows. The effect of scintillations on GPS time is explained. The effect of errors in the position of the antenna on the second method of time transfer is examined. A detailed examination of the performance achieved with the two methods is then presented, followed by discussions of two new types of GPS receiver systems. The paper concludes with a look at planned future enhancements to GPS. However, this paper is not just a review. It also includes important results from experiments performed at the National Physical Laboratory, New Delhi, India, on time-transfer precision, and the effects of factors such as scintillation.

The ICG

Appropriately, this issue also contains a report from URSI Vice President Gert Brussaard on the International Committee on Global Navigation Satellite Systems (ICG). This committee has its secretariat through the UN Office for Outer Space Affairs. The ICG met for the first time in November 2006, and URSI is involved in the committee as an Observer. The report includes an update on the committee's activities, a description of its organization, and information that should be very helpful to any URSI radio scientist who would like to become involved.

XXIXth General Assembly of URSI

The XXIXth General Assembly of URSI was held August 9-16, 2008, in Chicago, Illinois, USA. It was a wonderful meeting! Special thanks go to George Uslenghi, Danilo Erricolo, Sharad Laxpati, and all of the other members of the local Organizing Committee, and to Dr. M. K. Goel, the Scientific Program Coordinator.

Final production of this issue of the *Radio Science Bulletin* was delayed until after the General Assembly, and thus this modified version of my comments were able to be included. More news from the General Assembly will be included in the September issue.

URSI Forum on Radio Science and Telecommunications

Applications implemented through radio communication are exploding. This trend is likely to accelerate and, accordingly, the operation of many passive and active radio services will become dependent on the reliability of the new communication systems. All scientists engaged in radio-science activities must be aware of potential challenges and novel developments, and should be in a position to express their interest or/and concern. It is the objective of the "Forum on Radio Science and Telecommunications," organized on August 15 during the 2008 URSI GA in Chicago, to provide information, and to encourage exchanges on all domains of radio science and their relation to telecommunications.

The Forum will address the potential and challenges of novel developments in wireless telecommunications, which are of prime importance to anyone involved in radio science and system development, both from a user's and a policy maker's standpoint. It will have four invited papers of 40 minutes (25 minutes presentation plus 15 minutes discussion time) each, followed by a panel discussion. The papers will aim at stimulating discussions. The panel discussion of 40 minutes will focus on possible activities/ actions for URSI to deploy in this area.

The following subjects will be presented and discussed:

1. "Cognitive Radio" by Hiroshi Harada (JP)

Wireless systems in which the networks and/or the radio stations are sufficiently intelligent to provide and/or use ("grab") radio resources flexibly and adaptively;

2. "Ultra-wideband Wireless Systems" by Andy Molisch (USA)

Systems employing ultra-large bandwidth and/or signals of ultra-low spectral density; their unique features, potentials, and weaknesses;

3. "Interference Management" by Lilian Jeanty (NL)

Novel methods to protect communication and remote-sensing system integrity and to regulate spectrum usage in a dynamic and flexible manner, while catering to ultra-wideband and intelligent radio systems;

4. Health aspects of the proliferation of wireless communications in a deregulated arena by Niels Kuster (CH).

For the panel discussion, brief interventions (two minutes) may be prepared to complete the information or to raise specific points, such as:

- Other new trends in radio-system development
- Specific issues for the URSI Commissions
- New radio-science research areas to be developed in relation to new wireless services
- Possible URSI contributions to international discussions, the development of standards and regulations

The Forum discussions should result in the definition of follow-up actions (e.g., by one or more URSI working groups) to establish a more permanent role for URSI in representing the interests of radio science in worldwide discussions on the development of wireless communications.

Gert Brussaard
Radicom Consultants
Hendrik van Herenthalslaan 11
5737 ED Lieshout, Netherlands
Tel: +31 499 425430
E-mail: gert.brussaard@radicom.nl

Receiving the *Radio Science Bulletin* for the Next Triennium

The *Radio Science Bulletin* has historically been provided in print form by mail to all URSI Radioscientists. URSI Radioscientists are those who have registered for the most-recent past URSI General Assembly (the registration fee includes 40€ as the mandatory fee for becoming an URSI Radioscientist for the triennium), or those who apply to the URSI Secretariat to become URSI Radioscientists (and pay the fee) between General Assemblies. The *Bulletin* has also been sent in print form by mail to URSI Member Committees, and some libraries.

The costs of printing and mailing the *Bulletin* have increased dramatically over the past triennium, and the Web has become a viable alternative method of disseminating publications much more rapidly than is possible using ordinary mail. The version of the *Bulletin* available via the Web is also in color, in contrast to the printed version. In the interest of controlling costs and being able to continue to provide the *Radio Science Bulletin* to all URSI Radioscientists, at its May meeting the URSI Board decided to adopt a new policy for delivering the *Bulletin* in the coming triennium. Beginning with the XXIXth General Assembly in Chicago, 40€ will continue to be included as a required part of the registration fee to become an URSI Radioscientist (this is the approximate current cost of producing and publishing the *Bulletin* for the triennium, without the cost of printing and mailing individual copies). Starting with the March 2009 issue, all URSI Radioscientists will receive an e-mail alert when the *Bulletin* is available for download from the URSI Web site. If an URSI Radioscientist wishes to continue to receive a printed copy of the *Bulletin* by mail, he or she may pay an additional optional fee of 60€ to cover the printing and mailing costs for the triennium. Printed copies of the *Bulletin* will continue to be sent to Member Committees and libraries.

URSI Accounts 2007

Income in 2007 was higher than average due to the fact that some member committees paid their arrears and some paid their dues in advance. We also received the remainder of the fees of the GA2005 from New Delhi. Total income from the 2005 GA was approx. EUR 81,500 which is far less than the total cost of the GA to URSI which was around EUR 140,000. In real market value in EUR the investments have gained value even when taking into account the drop of the USD. The net total URSI assets have increased considerably. However, a significant part of these assets are allocated. The allocations include an important provision for the 2008 GA, leaving a reserve of less than EUR 100,000.

On the expenditure side the following points are noteworthy:

- The cost of the Coordinating Committee meeting: EUR 18,929.09
- The expenditure to the Commissions: EUR 33,897.50
- The travel and representation costs of the Board: EUR 17,091.54

The latter point reflects the increased efforts by the Board in representing URSI interests in international organisations. In the budget for the next triennium, therefore, a new item: "Travel by the Board" will be added. Also, provisions will be made for the Regional Centre in New Delhi. The Board is seriously considering means to limit the loss on future General Assemblies. There are three possibilities to accomplish this: reduce the travel cost of URSI officials, increase the revenue, or reduce the cost of the General Assembly by requiring less infrastructural demands for organizing it.

The General Secretariat has handled the day-to-day finances in its usual very efficient way, for which our sincere thanks.

G. Brussaard,
Treasurer, URSI

BALANCE SHEET: 31 DECEMBER 2007

ASSETS	EURO	EURO
Dollars		
Merrill Lynch WCMA	574.20	
Fortis	1,310.50	
Smith Barney Shearson	5,509.55	
		7,394.25
Euros		
Banque Degroof	2,483.33	
Fortis	52,939.57	
		55,422.90
Investments		
Demeter Sicav Shares	22,681.79	
Rorento Units	111,414.88	
Aqua Sicav	63,785.56	
Merrill-Lynch Low Duration (304 units)	3,268.17	
Massachusetts Investor Fund	250,483.18	
Provision for (not realised) less value	(7,472.37)	
Provision for (not realised) currency differences	(80,263.20)	
	363,898.01	
684 Rorento units on behalf of van der Pol Fund	12,414.34	
		376,312.35
Short Term Deposito		201,139.20
Petty Cash		942.15
Total Assets		641,210.85
Less Creditors		
IUCAF	10,778.95	
ISES	11,110.14	
		(21,889.09)
Balthasar van der Pol Medal Fund		(12,414.34)
NET TOTAL OF URSI ASSETS		**606,907.42**

The net URSI Assets are represented by:	EURO	EURO
Closure of Secretariat		
Provision for Closure of Secretariat		90,000.00
Scientific Activities Fund		
Scientific Activities in 2008	45,000.00	
Publications in 2008	40,000.00	
Young Scientists in 2008	40,000.00	
Administration Fund in 2008	85,000.00	
Scientific Paper submission software in 2008	30,000.00	
I.C.S.U. Dues in 2008	3,600.00	
		243,600.00
XXIX General Assembly 2008 Fund:		
During 2006-2007-2008		180,000.00
Total allocated URSI Assets		513,600.00
Unallocated Reserve Fund		93,307.42
		606,907.42

Statement of Income and expenditure for the year ended 31 December 2007

I. INCOME	EURO	EURO
Grant from ICSU Fund and US National		
Academy of Sciences	0.00	
Allocation from UNESCO to ISCU Grants Programme	0.00	
UNESCO Contracts	0.00	
Contributions from National Members (year -1)	33,578.29	
Contributions from National Members (year)	170,577.70	
Contributions from National Members (year +1)	38,789.50	
Contributions from Other Members	0.00	
Special Contributions	0.00	
Contracts	0.00	
Sales of Publications, Royalties	0.00	
Sales of scientific materials	0.00	
Bank Interest	3,534.04	
Other Income	46,173.50	
Total Income		**292,653.03**

II. EXPENDITURE

A1)	Scientific Activities		54,026.74
	General Assembly 2005/2008	18,929.09	
	Scientific meetings: symposia/colloqiua	33,897.50	
	Working groups/Training courses	0.00	
	Representation at scientific meetings	1,200.15	
	Data Gather/Processing	0.00	
	Research Projects	0.00	
	Grants to Individuals/Organisations	0.00	
	Other	0.00	
	Loss covered by UNESCO Contracts	0.00	

	EURO	EURO
A2) Routine Meetings		5,372.04
Bureau/Executive committee	5,372.04	
Other	0.00	
A3) Publications		34,266.58
B) Other Activities		10,682.00
Contribution to ICSU	4,682.00	
Contribution to other ICSU bodies	6,000.00	
Activities covered by UNESCO Contracts	0.00	
C) Administrative Expenses		74,350.69
Salaries, Related Charges	61,956.64	
General Office Expenses	3,313.14	
Travel and representation	17,091.54	
Office Equipment	3,081.30	
Accountancy/Audit Fees	5,187.88	
Bank Charges/Taxes	2,555.02	
Loss on Investments (realised/unrealised)	(18,834.83)	
Total Expenditure:		**178,698.05**
Excess of Income over Expenditure		113,954.98
Currency translation difference (USD => EURO) - Bank Accounts		(786.88)
Currency translation difference (USD => EURO) - Investments		(18,461.57)
Currency translation difference (USD => EURO) - others		0.00
Accumulated Balance at 1 January 2007		512,200.89
		606,907.42

Rates of exchange:

January 1, 2007	$ 1 = 0.7590 EUR
December 31, 2007	$ 1 = 0.6860 EUR

	EURO
Balthasar van der Pol Fund	
684 Rorento Shares: market value on December 31	
(Aquisition Value: USD 12.476,17/EUR 12.414,34)	29,815.56
Market Value of investments on December 31, 2006/2005	
Demeter Sicav	60,201.90
Rorento Units (1)	566,670.00
Aqua-Sicav	85,579.79
M-L Low Duration	2,070.84
Massachusetts Investor Fund	163,944.94
	878,467.48

(1) Including the 684 Rorento Shares of the van der Pol Fund

APPENDIX: Detail of Income and Expenditure

	EURO	EURO
I. INCOME		
Other Income		
Income General Assembly 2002	0.00	
Income General Assembly 2005	46,173.50	
Revenu Taxes	0.00	
		46,173.50
II . EXPENDITURE		
General Assembly 2005		
Organisation	0.00	
Vanderpol Medal	0.00	
Expenses officials	0.00	
Young scientists	0.00	
		0.00
Symposia/Colloquia/Working Groups:		
Commission A	0.00	
Commission B	9,000.00	
Commission C	3,000.00	
Commission D	2,500.00	
Commission E	0.00	
Commission F	3,500.00	
Commission G	3,000.00	
Commission H	2,000.00	
Commission J	0.00	
Commission K	9,000.00	
Central Fund	1,897.50	
		33,897.50
Contribution to other ICSU bodies		
UNESCO-ICTP	0.00	
FAGS	4,000.00	
IUCAF	2,000.00	
		6,000.00
Publications:		
Printing 'The Radio Science Bulletin'	12,708.78	
Mailing 'The Radio Science Bulletin'	21,557.80	
URSI Leaflet	0.00	
		34,266.58

URSI Awards 2008

The URSI Board of Officers decided at their May 2008 meeting in Ghent to follow the recommendations of the Awards Panel and to give the 2008 Awards to the following distinguished scientists:

The Balthasar Van der Pol Gold Medal

The Balthasar Van der Pol Gold Medal was awarded to Prof. William J. Welch with te citation:

"Pioneer of millimeter wavelenght interferometry to investigate astronomical objects ranging from solar system planets to galaxies at the edge of the Universe with spectral and angular resolution"

John Howard Dellinger Gold Medal

The John Howard Dellinger Gold Medal was be awarded to Prof. Alan Ernest E. Rogers with te citation:

"For his outstanding contributions to instrumentation in radio astronomy and its use to make fundamental discoveries about interstellar masers, superluminal expansion of quasars, deuterium abundance in the galaxy, and plate tectonics"

Appleton Prize

After considering the views submitted by the Awards Advisory Panel, the Board of Officers submitted a short list of candidates in order of preference, with reasons for the order, to the Royal Society.

The Council of the Royal Society approved the recommendation of the URSI Board to award the 2008 Appleton Prize to Prof. U.S. INAN with the citation:

For fundamental contributions to understanding of whistler-mode wave-particle interaction in near-Earth space and the electrodynamic coupling between lightning discharges and the upper atmosphere

Booker Gold Medal

The Booker Gold Medal was awarded to Prof. H. Matsumoto with te citation:

"For his outstanding contributions to the understanding of non-linear plasma wave processes, promotion of computer simulations in space plasma physics and international leadership in plasma wave research"

Koga Gold Medal

The Koga Gold Medal was awarded to Dr. D.F. Sievepiper, with te citation:

"For contributions to the development of artificial impedance surfaces and conformal antennas"

The Awards were presented to the Awardees during the Opening Ceremony of the XXIXth General Assembly in the Grand Ballroom of the Hyatt Regency Chicago Hotel, United States, on 10 August 2008

URSI and the International Committee on Global Navigation Satellite Systems (ICG): An Update

1. Background

Since its inaugural meeting in November 2006, URSI has been participating in the International Committee on Global Navigation Satellite Systems (ICG) as an observer. I reported on the early activities in the *Radio Science Bulletin* of March 2007.

The following is a short summary of the background. The "Vienna Declaration on Space and Human Development" of the United Nations calls for universal access to and compatibility of space-based navigation and positioning systems. In 2004, in Resolution 59/2, the UN General Assembly invited the system and service providers to consider establishing an international committee on GNSS, in order to maximize the benefits of the use and application of GNSS to support sustainable development. The ICG was established and held its inaugural meeting in Vienna in November 2006.

The secretariat is provided by the UN Office for Outer Space Affairs (OOSA) in Vienna. An Ad Hoc Working Group meets twice yearly to discuss current issues, and to prepare the annual Plenary Meetings of the ICG.

The ICG Terms of Reference define three types of participants: Members, Associate Members, and Observers. Associate Members and Observers both may advise the Committee and actively contribute to its work, but do not have a vote. There is very little difference between Associate Members and Observers: the latter will not be called upon to contribute financially. Among others, COSPAR, the ITU, and URSI participate as Observers. URSI's interests range from the development of propagation models and signal processing for radio navigation and positioning, to the application of GNSS in radio-science research, and hence cover the work areas of several Commissions.

At the first Plenary Meeting of the ICG in November 2006, the work plan for the ICG was adopted. This plan defines four working groups:

A. Compatibility and interoperability
B. Enhancement of performance of GNSS services
C. Information dissemination
D. Interaction with national and regional authorities and relevant international organizations, particularly in developing countries

The Annex contains a summary of the work plan of the Working Groups.

Furthermore, a Providers' Forum (with exclusive participation of providers of navigation satellite systems) was established, with the task to enhance compatibility and interoperability among current and future systems.

2. Second PlenaryMeeting of the ICG (Bangalore, September 5-7, 2007)

2.1 Providers' Forum

Immediately preceding the second Plenary Meeting, a meeting of the Providers' Forum was held on September 4. The agenda included system and service updates from each of the following providers:

- China: Compass/BeiDou Navigation Satellite System (CNSS)
- India: Global Positioning System and Geostationary (GEO) Augmented Navigation System (GAGAN) and Indian Regional Navigation Satellite System (IRNSS)
- Japan: Quasi-Zenith Satellite System (QZSS) and Multifunctional Transport Satellite (MTSAT) Satellite-based Augmentation System (MSAS);
- Russian Federation: Global Navigation Satellite System (GLONASS) and Wide-area System of Differential Corrections and Monitoring (SDCM);
- United States: Global Positioning System (GPS) and Wide-Area Augmentation System (WAAS);
- European Community: European Satellite Navigation System (Galileo) and European Geostationary Navigation Overlay Service (EGNOS).

Key issues discussed included spectrum protection, compatibility and interoperability, orbital debris de-confliction, availability of a free and open service, timely information dissemination, etc.

All system providers agreed that as a minimum, all GNSS signals and services must be compatible. To the maximum extent possible, open signals and services should also be interoperable. In this context, the following general definitions were agreed upon:

- *Compatibility* refers to the ability of space-based positioning, navigation, and timing services to be used separately or together without interfering with each individual service or signal;
- *Interoperability* refers to the ability of open global and regional satellite navigation and timing services to be used together to provide better capabilities at the user level than would be achieved by relying solely on one service or signal.

2.2 Experts Meeting

On September 5, an Experts Meeting was held, consisting of five scientific and technical sessions:

- Prediction of natural disaster and research on climate change and Earth science
- Geodetic reference frames
- Atomic time standards, Coordinated Universal Time (UTC), and time transfer
- Ionospheric/tropospheric models and space-weather effects
- GNSS activities in India

2.3 Second Plenary Meeting

The second Plenary Meeting was held on September 6 and 7. Reports from the Providers Forum and the Working Groups were presented, and the Committee considered the recommendations and plans for future work of the Working Groups. A special report was discussed on the subject of Satellite-based Augmentation System certification.

The Committee agreed that China be recognized as a provider of GNSS, since it was developing the Compass Navigation Satellite System.

Since this was only the second full meeting of the ICG, several actions were proposed to streamline the working methods of the Working Groups and to avoid duplications.

The Committee expressed its satisfaction with the continuing effort by the secretariat to develop the information portal of the Committee as part of the Web site of the Office for Outer Space Affairs.

The Committee accepted the invitation of the United States to host the third meeting, to be held in 2008, and noted the offer of the Russian Federation to host the fourth meeting in 2009.

Two informal meetings of the Ad Hoc Working Group to prepare the next meeting of the Committee were planned for 2008.

3. Conclusions

All presentations from the Providers' Forum and the technical expert sessions as well as the Working Group Reports are available at the portal of the Committee:

http://www.unoosa.org/oosa/en/SAP/gnss/icg.html

After a slow start, the work by the different Working Groups now seems to be developing. From the subjects of the technical sessions mentioned above, one can conclude that there are several areas where URSI is able to make contributions. It is planned to prepare a report for the next Plenary meeting, summarizing URSI's interest in various activities. These include both research and development concerning GNSS itself, and scientific applications of the use of the positioning and navigation system. Inputs to this report will be solicited from the Commissions involved.

4. Annex

4.1 Summary of the ICG Work Plan and Working Group Actions

The ICG work plan is executed by five Working Groups, which have defined a number of specific Actions.

4.1.1 Working Group A: Compatibility and Interoperability

The working group formed to address compatibility and interoperability, to be co-led by the United States of America and the Russian Federation, will pursue the following actions:

- *Action A1:* Establish a providers forum to enhance compatibility and interoperability among current and future global and regional space-based systems.
- *Action A2*: Organize a workshop(s) on measures being taken by Members, Associate Members, and Observers to enhance interoperability and compatibility of (1) global and regional space-based systems and (2) regional ground-based DGNSS.
- *Action A3*: Survey the level of interoperability and standardization among GNSS constellations and augmentations in order to identify concrete steps that can be taken at different levels (regulatory, system implementation, user algorithms) to improve interoperability and standardization. It is expected that the situation is well advanced in civil aviation and maritime areas, the effort would therefore probably need to concentrate on land-based applications and users.
- *Action A4:* Consider guidelines for the broadcast of natural disaster alarms via GNSS.

- *Action A5:* Develop a strategy for support by the International Committee of mechanisms to detect and mitigate sources of electromagnetic interference, taking existing regulatory mechanisms into consideration.

4.1.2 Working Group B: Enhancement of Performance of GNSS Services

As a unique combination of GNSS service providers and major user groups, the Committee will work to promote and coordinate activities aimed at enhancing GNSS performance, recommending system enhancements, and meeting future user needs. Specifically, the following actions will be taken by a working group co-led by India and the European Space Agency:

- *Action B1:* Develop a reference document on models and algorithms for ionospheric and tropospheric corrections.
- *Action B2:* Examine the problem of multipath and related mitigation actions affecting both GNSS systems and user receivers, especially for mobile receivers.
- *Action B3*: Examine the extension of GNSS service to indoor applications.

4.1.3 Working Group C: Information Dissemination

The Committee will consider the establishment of user information centers by GNSS providers. The maintenance of a globally focused Web site will be a major task of these centers. The United Nations, through the Office for Outer Space Affairs of the Secretariat and on behalf of the Committee, will combine all the Web sites into a single site to act as a portal for users of GNSS services. Therefore, the Office for Outer Space Affairs will lead a working group to accomplish the following actions:

- *Action C1:* Establish the International Committee information portal, drawing on contributions from Members, Associate Members, and Observers of the Committee. This will include a calendar of GNSS-related events.
- *Action C2:* Identify undergraduate and graduate courses on GNSS to be included on the Committee portal.
- *Action C3:* Consider the possibility of disseminating a list of relevant textbooks on GNSS in English and other languages through the Committee portal. Consideration will also be given to developing a glossary of terms and definitions.
- *Action C4:* Consider the use of the Regional Centres for Space Science and Technology Education, affiliated to the United Nations, to promote GNSS use and applications.
- *Action C5:* Identify international conferences where Members, Associate Members, and Observers will make presentations on the existence and work of the International Committee. A list of such events will be maintained on the Committee information portal.
- *Action C6:* Develop a proposal for further mechanisms to promote the applications of GNSS.

4.1.4 Working Group D: Interaction with National and Regional Authorities and Relevant International Organizations

The Committee will establish links with national and regional authorities and relevant international organizations, particularly in developing countries. The Committee will organize and sponsor regional workshops and other types of activity in order to fulfill its objectives. The Fédération Internationale des Géomètres (FIG), the International Association of Geodesy (IAG), and the International GNSS Service will co-lead the activities listed below:

- *Action D1:* Define minimum operational performance standards for GNSS performance monitoring networks.
- *Action D2:* Establish a working group focused on site quality, integrity, and interference monitoring (SQII).
- *Action D3:* Establish a working group to develop a strategy for support by the International Committee of regional reference systems (e.g., the African Geodetic Reference Framework (AFREF), the European Position Determination System (EUPOS), the IAG Reference Frame Sub-Commission for Europe (EUREF), and the Geocentric Reference System for the Americas (SIRGAS)).
- *Action D4:* Establish a working group to develop a strategy for support by the International Committee of mechanisms to detect and mitigate sources of electromagnetic interference, taking existing regulatory mechanisms into consideration.

4.1.5 Working Group E: Coordination

In the future, the Committee will consider, make recommendations, and agree on actions to promote appropriate coordination across GNSS programs. Furthermore, the Committee will encourage its Members, Associate Members, and Observers to maintain communication, as appropriate, with other groups and organizations involved in GNSS activities and applications, through the relevant channels within their respective governments and organizations. The Committee could also support the establishment of national and/or regional planning groups for GNSS that would address regulations associated with the use of GNSS services, and suggest organizational models to use at the national level for co-coordinating and governing GNSS use.

Gert Brussaard
Radicom Consultants
Hendrik van Herenthalslaan 11
5737 ED Lieshout, Netherlands
Tel: +31 499 425430
E-mail: gert.brussaard@radicom.nl

Introduction to Special Sections Honoring Jenifer Haselgrove

Even a significant historical achievement may diminish in its apparent magnitude if too casually viewed through the prism of our present state. This is vividly portrayed in the world of computing, where an information age now looks back on enormous machines, with their clumsy human interfaces and relatively small processing and memory capabilities. Of course, more careful scrutiny reveals that what was achieved with those relatively primitive computers was far from trivial. After all, the birth of computers was concurrent with the second World War, and success and failure could be seen in very stark terms. The first automated computers emerged in Germany, the US, and the UK in the 1940s. EDSAC, the first UK general-purpose stored-program computer, was commissioned at Cambridge in 1949.

In 1950s' Cambridge, the scene was set for a young PhD student named Jenifer Haselgrove (Figure 1) to make a remarkable connection, which would end up greatly benefiting the world's radio-science community.

Figure 1. Jenifer Haselgrove

According to Haselgrove, "these were heady days, where everything you did on a computer was new." An abstract from her recent musings on those times provides insight, and it is curious to get a glimpse of the early utilization of computer-science constructs that have stood the test of time, such as functions and libraries:

> The programming language was a simple mnemonic code, for example, "A 150" meant "add the number in storage location 150 into the accumulator [the place where the current working number was]." You didn't have to think in binary arithmetic, contrary to what you read in some historical accounts of early computers nowadays. The program (that spelling was settled on for use even in Britain fairly early, after some argument) was typed on to 5-hole (teleprinter) paper tape and read into the machine by a single built-in program of "initial orders." The instructions making up programs were called orders, which may have been a bit confusing at first, but nothing like the terrible mistake made in FORTRAN and later languages, where instructions are called "statements." Of course, programs didn't all do numerical work – take the initial orders for example – and there had to be programs right from the beginning to read users' programs and data and to print results. Those were part of a whole "library" of subroutines for standard jobs, including numerical ones such as calculating square roots, and even division, which wasn't originally built in. At first the output was directly on to a teleprinter, typing on paper at, I think, 7 characters a second, but later that was changed to punched tape like the input, to be printed later or while the machine got on with the calculating. The actual computing building was the original Anatomy Department, and there was a coffin lift (elevator) in one corner. I don't think any of us bothered about that, though, even working alone in the building in the middle of the night. We all had to do some of that – machine time was rare and valuable!

Jenifer Haselgrove had begun her tertiary education after moving to Cambridge (Figure 2) in 1948, where she started reading math. At that time, she connected with Brian Haselgrove when they both joined a music group, which listened to "long playing records," a recording and playback technology that intrigued them both. They married in 1951. In the same year, she moved to physics and the Cavendish Laboratory. When meeting the seniors on the board of the Cavendish Laboratory, Haselgrove was a little taken back to find them expressing relief that she was not interested in practical or field work. Curious on whether she was running up against some Cavendish misogyny, she was amused to hear later, after some discrete inquiry with the departmental secretary, that the panel was greatly concerned about the lack of "facilities" at the field station! Furthermore, the secretary had chided the board for being silly, suggesting "they could use one bush and Haselgrove another."

Figure 2. Richard Haselgrove, Jenifer Haselgrove's brother, supplied this photograph from earlier Cambridge days. Jenifer is fourth from right, front row. Brian Haselgrove (her first husband) is fifth from right, front row. John Leech (her second husband) is fifth from left, back row.

Haselgrove was just starting her research when she fell pregnant, and 1952/3 was spent preparing for and rearing her son. Her early postgraduate research was performed under the renowned radio scientist Kenneth Budden, where she started on radio propagation, spending time calculating the trajectory of short ray paths by hand. Many other researchers were similarly developing mathematical and analytic techniques and constructs to further the domain of radio ray tracing through that time. Shortly afterwards, in discussions with her husband – himself a highly respected mathematician who was now working with EDSAC in the math department – she was encouraged to look at ways to solve the ray-propagation problem on a computer. Her first work on the topic was published shortly afterwards, and over the next five years a series of papers were published (all referenced in the articles to follow). Subsequent to her breakthroughs, there is record of numerous efforts in computer-based ray tracing emerging throughout the world. Jenifer Haselgrove had developed the connection that resulted in the birth of her second creation, computer-based radio ray tracing. By the early 1960s, the radio-science community was on board and was enamored with her latest "child." Through popular use, her name is now enshrined with the original formulations, viz., the *Haselgrove Equations*. The equations derive from the physical-principle formulations of William Rowan Hamilton and James Clerk Maxwell, when applied to radio propagation in the ionosphere. Haselgrove herself referred to them as Hamilton's ray equations.

Haselgrove worked on ray-tracing problems in collaboration with her husband, until he died in 1964. After that, she remarried to John Leech, another Cambridge mathematician who had worked and socialized with the Haselgroves through their careers. Jenifer Haselgrove worked as a computer scientist at Glasgow University, and retired in 1981.

A two-part collection of articles in the June and December editions of the *Radio Science Bulletin* has been developed. It amounts to a dedicated effort by a group of radio scientists who have benefited from the path pioneered by Jenifer Haselgrove, and have taken the time to honor that achievement.

Of note is the variety of applications to which the Haselgrove Equations and computer-based ray tracing have been applied. Topics include EHF, VLF, and HF electromagnetic propagation in the magnetosphere; whistlers; HF communications; inter-satellite communications; HF radar; HF radio direction finding; and remote-sensing investigations of the ionosphere and magnetosphere. Many of these topics are touched on with the following papers. The same equations also provide solutions for acoustic ray tracing.

Part 1 of the testimonial editions appears in this issue. It starts with a theoretical treatise by Coleman, who demonstrates a method of deriving the Haselgrove Equations directly from Maxwell's equations. This is followed by another theoretical developmental piece by Walker, who demonstrates a derivation based on Fermat's principle. Walker goes on to show some historical applications of Haselgrove's methods in modeling electromagnetic propagation in the magnetosphere.

Nickich discusses his career and the foundational role that radio ray tracing, based on the Haselgrove Equations, played in this. Finally, Barnes discusses two applications of ray tracing at HF, pertaining to sounders and radar.

In the December 2008 issue, Part 2 of the testimonial editions will appear. This will begin with a summary of multiple applications of the Haselgrove Equations by James. It will be followed by a historical summary of ray-tracing development by Bennett. Dyson discusses the impact of the Haselgrove Equations and ray tracing on his career in radio science.

Kimaru discusses a career that began almost simultaneously with Haselgrove's, which involved the application of her ray-tracing techniques to whistler and other wave propagation in the magnetosphere. The edition is completed with a compilation of applications by Bertel.

Please enjoy this tribute to Jenifer Haselgrove and her achievements. In collating and researching for this work, we have been inspired by many of the great minds that have brought us so far in radio science, but none more than Jenifer Haselgrove and those around her at Cambridge, which led to her famous formulation and her use of one of mankind's first computers to solve it. Our hope is that these editions draw your attention to those "heady days," a person who lived through them, and her achievements. Just as importantly, it is hoped that it inspires you in your radio-science endeavors so that someone may write about them with great respect in the near future.

Rod Barnes and Phil Wilkinson
Co-Guest Editors,
Special Sections Honoring Jenifer Haselgrove
E-mail: rbarnes@rri-usa.org, phil@ips.gov.au

Ionospheric Ray-Tracing Equations and their Solution

Christopher J. Coleman

Abstract

The Haselgrove ray-tracing equations are derived directly from Maxwell's equations. Some methods for their solution are discussed. In particular, we look at reduced versions of the equations that allow fast numerical solution and, in some cases, analytic solution.

1. Introduction

It is now over 50 years since Jenifer Haselgrove introduced the ray-tracing equations that have become synonymous with her name. During this time, her equations have become a major tool for investigating radiowave propagation in the ionosphere. In the early days, such studies were needed because of the importance of ionospheric propagation for long-range terrestrial communications. Such communications take place at high frequencies (HF), and use the fact that radiowaves at HF (3 to 30 MHz) are refracted back down to the Earth by the ionosphere. However, with the advent of artificial satellites, ionospheric communication has become less important. Nevertheless, such communications are still an important tool for the military, aid agencies, and remote communities. Furthermore, the introduction of over-the-horizon radar (OTHR) has significantly increased the use of ionospheric propagation. The extreme demands of over-the-horizon radar have made it necessary to understand ionospheric propagation at a more refined level, and the Haselgrove equations have played an important part in such studies.

Starting with the Haselgrove equations, we review some of the ray-tracing approaches that are available for the study of ionospheric propagation. In Section 2, we derive the Haselgrove equations directly from Maxwell's equations (together with some basic plasma physics). Derivations of the Haselgrove equations tend to use ray optics, in particular the Hamiltonian equations, as their starting point [1-3]. Section 2 attempts to start from a more fundamental position, and to cast the equations as part of a procedure for the solution of Maxwell's equations in the high-frequency limit. In [4], it was found that the Runge-Kutta-Fehlberg numerical scheme constituted a very efficient means of solving the Haselgrove equations, and so we include a brief description of this algorithm. Unfortunately, there still exist ray-tracing applications for which computer solutions to the Haselgrove equations are not fast enough. A particular case is the coordinate registration (CR) problem of over-the-horizon radar. In the coordinate-registration problem, fast ray tracing is required to convert the radar range (the time for the radio signal to travel to the target) into the actual ground range. Due to the ever-changing nature of the ionosphere, these calculations need to be done in real time. Consequently, in Sections 3 to 5 we look at some simplifications to the Haselgrove equations that can provide this increased speed. In Section 3, we look at the situation where the background magnetic field can be regarded as being weak (a good approximation for most HF frequencies above 10 MHz). In Section 4, we look at the simplification that results when we totally ignore the background magnetic field, an approximation that can be made more respectable by the use of effective wave frequencies. Finally, in Section 5, we consider some first integrals of the ray-tracing equations, and we also consider analytic solutions that can be derived from these first integrals.

2. The Haselgrove Equations

For time-harmonic fields in a vacuum, Maxwell's equations yield the field equations

$$\nabla\times\nabla\times\underline{E}-\omega^2\mu_0\varepsilon_0\underline{E}=-j\omega\mu_0\underline{J}\,, \qquad (1)$$

where $\underline{E}$ is the time-harmonic electric field, $\underline{J}$ is the time-harmonic current density, and ω is the wave frequency. Within the ionospheric plasma, the motion of an electron satisfies

Christopher J. Coleman is with the School of Electrical and Electronic Engineering, University of Adelaide, Adelaide, SA 5005, Australia;
e-mail: ccoleman@eleceng.adelaide.edu.au

$$m\dot{\underline{w}} = e\underline{E} + e\underline{w}\times\underline{B}_0 - m\nu\,\underline{w}, \qquad (2)$$

where m is the electron mass, e is the electron charge, ν is the plasma collision frequency, $\underline{w}$ is the electron velocity, and $\underline{B}_0$ is the magnetic field of the Earth. The plasma current is given by $eN\underline{w}$, where N is the electron density, Thus, for a time-harmonic field, Equation (2) reduces to

$$-j\omega\varepsilon_0 X\,\underline{E} = U\,\underline{J} + j\underline{Y}\times\underline{J}\,, \qquad (3)$$

where $U = 1 - j\nu/\omega$, $X = \omega_p^2/\omega^2$, $\underline{Y} = -e\underline{B}_0/m\omega$, and $\omega_p = \sqrt{Ne^2/\varepsilon_0 m}$ is the plasma frequency. In the high-frequency limit ($\omega \to \infty$), we normally assume an electric field of the form $\underline{E} = \underline{E}_0 \exp(-j\beta\varphi)$, where $\underline{E}_0$ and φ are both slowly varying functions of position and $\beta = \omega\sqrt{\mu_0\varepsilon_0}$ (see [5] and [6] for more information concerning the high-frequency approximation). Noting that

$$\nabla\times\nabla\times\underline{E} = \nabla(\nabla\bullet\underline{E}) - \nabla^2\underline{E}\,,$$

and substituting the assumed form of solution into Equation (1), we find the leading order in ω yields

$$-\nabla\varphi(\nabla\varphi\bullet\underline{E}_0) + (\nabla\varphi\bullet\nabla\varphi)\underline{E}_0 - \underline{E}_0 = \frac{-j\omega\mu_0}{\beta^2}\underline{J}_0 \qquad (4)$$

where $\underline{J} = \underline{J}_0\exp(-j\beta\varphi)$. From Equation (3), it follows that

$$-j\omega\varepsilon_0 X\,\underline{E}_0 = U\,\underline{J}_0 + j\underline{Y}\times\underline{J}_0\,. \qquad (5)$$

The dot product of Equation (4) with $\nabla\varphi$ yields

$$\nabla\varphi\bullet\underline{E}_0 = \frac{j\omega\mu_0}{\beta^2}\nabla\varphi\bullet\underline{J}_0\,. \qquad (6)$$

Eliminating $\nabla\varphi\bullet\underline{E}_0$ from Equation (4) using Equation (6), we obtain

$$(\nabla\varphi\bullet\nabla\varphi - 1)\underline{E}_0 = \frac{-j\omega\mu_0}{\beta^2}\left(\underline{J}_0 - \nabla\varphi\nabla\varphi\bullet\underline{J}_0\right) \qquad (7)$$

Eliminating $\underline{E}_0$ using Equation (5),

$$(\nabla\varphi\bullet\nabla\varphi - 1)\left(U\,\underline{J}_0 + j\underline{Y}\times\underline{J}_0\right) = -X\left(\underline{J}_0 - \nabla\varphi\nabla\varphi\bullet\underline{J}_0\right) \qquad (8)$$

From the dot product of Equation (8) with $\nabla\varphi$, we obtain

$$-j\nabla\varphi\bullet\left(\underline{Y}\times\underline{J}_0\right) = -j(\nabla\varphi\times\underline{Y})\bullet\underline{J}_0 = (U - X)\nabla\varphi\bullet\underline{J}_0 \qquad (9)$$

Using this to eliminate $\nabla\varphi\bullet\underline{J}_0$ from Equation (8),

$$(U\nabla\varphi\bullet\nabla\varphi - U + X)\underline{J}_0 = \frac{-jX}{U - X}\nabla\varphi(\nabla\varphi\times\underline{Y})\bullet\underline{J}_0 - j\underline{Y}\times\underline{J}_0\,(\nabla\varphi\bullet\nabla\varphi - 1) \qquad (10)$$

Without loss of generality, we choose $Y_1 = Y$ and $Y_2 = Y_3 = 0$ (i.e., the x_1 coordinate direction is in the direction of the magnetic field). In this case, Equation (10) can be rewritten as

$$\begin{bmatrix} \left(Up^2 + X - U\right) & jYPp_1p_3 & -jYPp_1p_2 \\ 0 & \left(Up^2 + X - U\right) + jYPp_2p_3 & -jYPp_2^2 - jY\left(p^2 - 1\right) \\ 0 & jYPp_3^2 + jY\left(p^2 - 1\right) & \left(Up^2 + X - U\right) - jYPp_3p_2 \end{bmatrix}\underline{J}_0 = \underline{0} \qquad (11)$$

where $\underline{p} = \nabla\varphi$, $p^2 = \underline{p}\bullet\underline{p}$, and $P = X/(U - X)$. This will only have a nontrivial solution if the determinant of the matrix is zero, that is,

$$\left(Up^2 + X - U\right)\left[\left(Up^2 + X - U\right)^2 + Y^2P^2p_3^2p_2^2 - Y^2\left(Pp_3^2 + p^2 - 1\right)\left(Pp_2^2 + p^2 - 1\right)\right] = 0\,, \qquad (12)$$

from which either $Up^2 + X - U = 0$ (in which case, $\underline{E}_0$ is parallel to $\underline{Y}$), or

$$\left(Up^2 + X - U\right)^2 - Y^2\left(p^2 - 1\right)^2 - Y^2P\left(p^2 - 1\right)\left(p^2 - p_1^2\right) = 0\,, \qquad (13)$$

for which $\underline{E}_0$ has no component in the direction of $\underline{Y}$. Returning to a general system of coordinates (NB: $Yp_1 = \underline{Y}\bullet\underline{p}$ and $Y^2 = \underline{Y}\bullet\underline{Y}$), we obtain from Equation (13) that

$$\left(Up^2 + X - U\right)^2 - Y^2\left(p^2 - 1\right)^2 - P\left(p^2 - 1\right)\left[Y^2p^2 - \left(\underline{p}\bullet\underline{Y}\right)^2\right] = 0\,. \qquad (14)$$

Equation (14) is a first-order partial differential equation for the phase field, φ, and this can be solved by standard methods for partial differential equations. For the equation

$F(\underline{x}, \underline{p}) = 0$, we can define a solution in terms of characteristic curves (our rays). These, together with the values of φ along them, will satisfy the generalized Charpit equations (see [7], for example):

$$\frac{dx_i}{\partial F/\partial p_i} = \frac{d\varphi}{\sum_{j=1}^{3} p_j \, \partial F/\partial p_j} = \frac{dp_i}{-\partial F/\partial x_i} \cdot \qquad (15)$$

The above equations can be used to find the ray curves in terms of the values of the phase factor, φ, along them. We will consider the collision-free case ($U = 1$). Equation (14) then implies

$$\left(p^2 - 1\right)^2 \left(1 - Y^2 - PY^2\right) + 2\left(p^2 - 1\right)X + X^2$$

$$-P\left(p^2 - 1\right)Y^2 + P\left(p^2 - 1\right)\left(\underline{Y} \bullet \underline{p}\right)^2 = 0 \cdot \qquad (16)$$

Multiplying through by $1 - X$, Equation (16) can be recast as $F\left(\underline{x}, \underline{p}\right) = 0$, where

$$F\left(\underline{x}, \underline{p}\right) = \left(p^2 - 1\right)^2 \left(1 - X - Y^2\right)$$

$$+\left(p^2 - 1\right)\left[2X\left(1 - X\right) - XY^2\right]$$

$$+X^2\left(1 - X\right) + X\left(p^2 - 1\right)\left(\underline{Y} \bullet \underline{p}\right)^2 \cdot \qquad (17)$$

Consequently,

$$\frac{\partial F}{\partial p_i}$$

$$= 4p_i\left(1 - X - Y^2\right)\left(p^2 - 1\right) + 2p_i\left[2X\left(1 - X\right) - XY^2\right]$$

$$+2p_i X\left(\underline{p} \bullet \underline{Y}\right)^2 + 2Y_i X\left(p^2 - 1\right)\underline{p} \bullet \underline{Y} \cdot \qquad (18)$$

We introduce a new parameter, t, along the characteristic curves defined by

$$dt = -Xd\varphi \Big/ \sum_{j=1}^{3} p_j \, \partial F/\partial p_j$$

and then Equation (15), together with a rearranged version of Equation (18), yields

$$\frac{dx_i}{dt} = -\frac{4p_i}{X}\left(1 - X - Y^2\right)\left(p^2 - 1 + X\right) - 2p_i Y^2$$

$$-2p_i\left(\underline{p} \bullet \underline{Y}\right)^2 - 2Y_i\left(p^2 - 1\right)\underline{p} \bullet \underline{Y} \cdot \qquad (19)$$

If we now introduce the new dependent variable $q = \left(p^2 - 1\right)\Big/X$ (as defined in [3]), we obtain

$$\frac{dx_i}{dt} = -4p_i\left(1 - X - Y^2\right)\left(q^2 + 1\right) - 2p_i Y^2$$

$$-2p_i\left(\underline{p} \bullet \underline{Y}\right)^2 - 2Y_i\left(p^2 - 1\right)\underline{p} \bullet \underline{Y}, \qquad (20)$$

which is the Haselgrove equation for the advancement of $\underline{p}$ [2]. The other equation implied by Equation (15) requires the coordinated derivatives of F, that is

$$\frac{\partial F}{\partial x_i} = -\left(p^2 - 1\right)^2\left(\frac{\partial X}{\partial x_i} + 2\underline{Y} \bullet \frac{\partial \underline{Y}}{\partial x_i}\right)$$

$$+\left(p^2 - 1\right)\left[\frac{\partial X}{\partial x_i}\left(2 - 4X - Y^2\right) - 2X\underline{Y} \bullet \frac{\partial \underline{Y}}{\partial x_i}\right]$$

$$+X\frac{\partial X}{\partial x_i}\left(2 - 3X\right) + \frac{\partial X}{\partial x_i}\left(p^2 - 1\right)\left(\underline{Y} \bullet \underline{p}\right)^2$$

$$+2\underline{p} \bullet \frac{\partial \underline{Y}}{\partial x_i} X\left(p^2 - 1\right)\left(\underline{Y} \bullet \underline{p}\right) \cdot \qquad (21)$$

Introducing the same dependent and independent variables as for Equation (20), we obtain from Equation (15) that

$$\frac{dp_i}{dt} = \left\{\left[-\left(p^2 - 1\right) + \left(\underline{p} \bullet \underline{Y}\right)^2 + 2\left(1 - 2X\right) - Y^2\right]q\right.$$

$$\left.+2 - 3X\right\}\frac{\partial X}{\partial x_i}$$

$$+\sum_{j=1}^{3}\left(p^2 - 1\right)\left[2\left(\underline{p} \bullet \underline{Y}\right)p_j - 2\left(q + 1\right)Y_j\right]\frac{\partial Y_j}{\partial x_i} \qquad (22)$$

Equations (20) and (22) together constitute the original

Haselgrove equations [2]. Furthermore, recasting Equation (16) in terms of q [2], we obtain

$$q^2\left(1-X-Y^2\right)+q\left[\left(\underline{Y}\bullet\underline{p}\right)^2+2\left(1-X\right)-Y^2\right]+1-X=0 \tag{23}$$

from which q can be derived (note that there are two possible values of q, corresponding to two of the possible solutions to Equation (11)). The corresponding current, and hence the electric field through Equation (5), can be obtained from the homogeneous Equations (11) using the calculated values of $\underline{p}$ on the ray. However, Equations (11) do not determine the magnitude of the field. This must be calculated by tracking the power out from the source. This can be done by tracing out ray tubes, and noting that within these tubes, the power, $Ap\left|\underline{E}_0\right|^2/\left(2\eta_0\right)$, must remain constant ($A$ is the cross-sectional area of the tube). The magnitude of the electric field can then be calculated from the power.

There remains the issue of solving the above ray-tracing equations. Starting with some initial value, $\underline{w}(0)$, for $\underline{w}$, the system of ordinary differential equations, $d\underline{w}/dt=\underline{W}\left(\underline{w}\right)$, can be solved using Runge-Kutta techniques. This was suggested in the Haselgrove and Haselgrove paper [2]. If we start a ray at the Earth's surface, the initial values of $\underline{p}$ are given by the direction cosines at this starting point. However, the extreme variations that can exist in the ionosphere often make it necessary to vary the increment in t at each stage of the solution process, in order to maintain accuracy. Haselgrove and Haselgrove [2] suggested the introduction of a scaling to produce this effect, but the Runge-Kutta-Fehlberg (RKF) method (see [8] for example) can also achieve this.

Consider the solution at time t and prospective increment Δt. Form the vectors $\underline{k}_1$, $\underline{k}_2$, $\underline{k}_3$, $\underline{k}_4$, $\underline{k}_5$, and $\underline{k}_6$ through the process

$$\underline{k}_1=\Delta t\underline{W}\left(\underline{w}\right), \tag{24}$$

$$\underline{k}_2=\Delta t\underline{W}\left(\underline{w}+\underline{k}_1/4\right), \tag{25}$$

$$\underline{k}_3=\Delta t\underline{W}\left(\underline{w}+3\underline{k}_1/32+9\underline{k}_2/32\right), \tag{26}$$

$$\underline{k}_4=\Delta t\underline{W}\left(\underline{w}+1932\underline{k}_1/2197-7200\underline{k}_2/2197+7296\underline{k}_3/2197\right), \tag{27}$$

$$\underline{k}_5=\Delta t\underline{W}\left(\underline{w}+439\underline{k}_1/216-8\underline{k}_2+3680\underline{k}_3/513-845\underline{k}_4/4104\right), \tag{28}$$

$$\underline{k}_6=\Delta t\underline{W}\left(\underline{w}-8\underline{k}_1/27+2\underline{k}_2-3544\underline{k}_3/2565+1859\underline{k}_4/4104-11\underline{k}_5/40\right). \tag{29}$$

The approximate solution, $\underline{w}(t)+\Delta\underline{w}$ at $t+\Delta t$, is obtained using

$$\Delta\underline{w}_4=25\underline{k}_1/216+1408\underline{k}_3/2565+2197\underline{k}_4/4104-\underline{k}_5/5 \tag{29}$$

for the fourth-order Runge-Kutta method, and by

$$\Delta\underline{w}_5=16\underline{k}_1/135+6656\underline{k}_3/12825+28561\underline{k}_4/56430-9\underline{k}_5/50+2\underline{k}_6/55 \tag{30}$$

for the fifth-order method. The global truncation error will be $O\left(\Delta t^4\right)$ for the fourth-order method and $O\left(\Delta t^5\right)$ for the fifth-order method. The magnitude of the difference between the fourth- ($\underline{w}_4$) and fifth- ($\underline{w}_5$) order estimates, $\Delta w=\left|\underline{w}_5-\underline{w}_4\right|$, gives an estimate of the global truncation error in the fourth-order method.

The Runge-Kutta-Fehlberg method proceeds by adjusting the step at each stage so that this global error remains close to a pre assigned value, ΔE. That is, at each stage we adjust Δt so that

$$\Delta t_{new}=\Delta t_{old}\left(\frac{\Delta E}{\Delta w}\right)^{\frac{1}{4}}. \tag{31}$$

This approach has been used in [4] to provide an efficient numerical implementation of the Haselgrove equations.

3. The Weak-Background-Field Limit

For the collision-free case ($U=1$), Equation (3) can be inverted to yield

$$\underline{J} = \frac{-j\omega\varepsilon_0 X}{1-Y^2}\left(\underline{E} - j\underline{Y}\times\underline{E} - \underline{Y}\bullet\underline{E}\underline{Y}\right)\cdot \qquad (32)$$

Up until now, we have assumed that both the plasma frequency, ω_p, and the gyro frequency, $\omega_H = \left|e\underline{B}_0\right|/m$, are of the same order as the wave frequency, ω. We will now assume that $\omega >> \omega_p$ and, hence, we can ignore the term of second order, Y^2, in Equation (32). Substituting the trial solution $\underline{E} = \underline{E}_0 \exp\left(-j\beta\varphi\right)$ into Equation (1), the β^2 and β terms yield

$$-\nabla\varphi\nabla\varphi\bullet\underline{E}_0 + \left(\nabla\varphi\bullet\nabla\varphi\right)\underline{E}_0 - \left(1-X\right)\underline{E}_0 = 0 \qquad (33)$$

and

$$-\nabla\varphi\nabla\bullet\underline{E}_0 - \nabla\left(\nabla\varphi\bullet\underline{E}_0\right) + 2\nabla\varphi\bullet\nabla\underline{E}_0 + \nabla^2\varphi\underline{E}_0$$

$$= \beta X\underline{Y}\times\underline{E}_0, \qquad (34)$$

respectively. By taking the divergence of Equation (1), we obtain $\nabla\bullet\underline{E} = \left(j/\omega\varepsilon_0\right)\nabla\bullet\underline{J}$, from which and $\nabla\varphi\bullet\underline{E}_0 = 0$

$$\left(1-X\right)\nabla\bullet\underline{E}_0 = \nabla X\bullet\underline{E}_0 - \beta X\nabla\varphi\bullet\left(\underline{Y}\times\underline{E}_0\right)$$

for the two leading orders. Consequently, Equations (33) and (34) reduce to

$$\left(\nabla\varphi\bullet\nabla\varphi - 1 + X\right)\underline{E}_0 = 0 \qquad (35)$$

and

$$\frac{-\nabla\varphi}{1-X}\nabla X\bullet\underline{E}_0 + \frac{\beta X}{1-X}\nabla\varphi\nabla\varphi\bullet\left(\underline{Y}\times\underline{E}_0\right)$$

$$+2\nabla\varphi\bullet\nabla\underline{E}_0 + \nabla^2\varphi\underline{E}_0 = \beta X\underline{Y}\times\underline{E}_0 \qquad (36)$$

From Equation (35), we obtain the ray-tracing equations $d\underline{p}/dt = \nabla n^2/2$ and $d\underline{x}/dt = \underline{p}$, where $dt = d\varphi/p^2$ (i.e., t is the group distance) and $n^2 = 1-X$. These equations do not contain the background magnetic field, and hence can be solved far more efficiently than the original Haselgrove equations. Once again, the Runge-Kutta-Fehlberg method can be used with great effect.

Equation (36) can now be recast as a differential equation for $\underline{E}_0$ along a ray:

$$2\frac{d\underline{E}_0}{dt} + \frac{d\underline{x}}{dt}\nabla\ln n^2\bullet\underline{E}_0 + \nabla^2\varphi\underline{E}_0$$

$$= \beta X\left[\underline{Y}\times\underline{E}_0 - \frac{1}{1-X}\frac{d\underline{x}}{dt}\frac{d\underline{x}}{dt}\bullet\left(\underline{Y}\times\underline{E}_0\right)\right]\cdot \qquad (37)$$

Taking the dot product of Equation (37) with $\underline{E}_0$, we find that $d\left|\underline{E}_0\right|^2/dt + \nabla^2\varphi\left|\underline{E}_0\right|^2 = 0$. Combining this with Equation (37), we obtain an equation for the polarization vector, $\underline{P} = \underline{E}_0/\left|\underline{E}_0\right|$, of the form

$$2\frac{d\underline{P}}{dt} + \frac{d\underline{x}}{dt}\nabla\ln n^2\bullet\underline{P}$$

$$= \beta X\left[\underline{Y}\times\underline{P} - \frac{1}{1-X}\frac{d\underline{x}}{dt}\frac{d\underline{x}}{dt}\bullet\left(\underline{Y}\times\underline{P}\right)\right] \qquad (38)$$

Essentially, the background magnetic field manifests itself as a rotation of the polarization vector about the ray direction.

4. The Negligible-Background-Field Limit

When the magnetic field of the Earth can be ignored ($\underline{B}_0 = 0$), the Haselgrove equations can be reduced to

$$\frac{d}{ds}\left(n\frac{d\underline{x}}{ds}\right) = \nabla n, \qquad (39)$$

where s is the distance along the ray path. The equations are themselves the Euler-Lagrange (EL) equations for the variational principle

$$\delta\int_T^R n\left(\underline{x}\right)ds = 0, \qquad (40)$$

i.e., Fermat's principle. If we now consider the case for a ray path that remains in a plane, we can use a polar coordinate system $\left(r,\theta\right)$, and the variational Euler-Lagrange principle becomes

$$\delta\int_T^R n\left(r,\theta\right)\left[\left(\frac{dr}{d\theta}\right)^2 + 1\right]^{1/2} d\theta = 0\cdot \qquad (41)$$

The above ray-tracing equations can be used when any deviations from the great-circle path (certainly the case for a uniform ionosphere) are negligible. Coordinate r is the distance from the center of the Earth, and θ is an angular coordinate along the great-circle path. The Euler-Lagrange equations for the above variational principle can be reduced [9] to

$$\frac{dQ}{d\theta} = \frac{1}{2}\frac{r}{\sqrt{n^2 - Q^2}}\frac{\partial n^2}{\partial r} + \sqrt{n^2 - Q^2} \qquad (42)$$

and

$$\frac{dr}{d\theta} = \frac{rQ}{\sqrt{n^2 - Q^2}}, \qquad (43)$$

where $Q = n(dr/d\theta)\Big/\sqrt{(dr/d\theta)^2 + r^2}$. We now only need to solve two ordinary differential equations to find a ray, a lot more efficient than the four equations for the planar Haselgrove equations. The downside is that we have ignored lateral deviations and the background magnetic field. Some measure of the background magnetic field can be incorporated by ray tracing at effective wave frequencies $\omega \pm \omega_H/2$, where ω_H is the gyro frequency at the midpoint of the ray path (the different effective frequencies correspond to the two different roots of Equation (23)). (The reader should refer to [10] for more accurate expressions of effective frequency.) Once again, the Runge-Kutta-Fehlberg technique can be used to efficiently solve the ray-tracing equations.

If we move to a nearby ray path, the deviations (δr and δQ) in quantities r and Q can be calculated from

$$\frac{d(\delta Q)}{d\theta} = \frac{r}{\sqrt{n^2 - Q^2}}\left[\frac{\delta r}{2}\left(\frac{1}{r}\frac{\partial n^2}{\partial r} + \frac{\partial^2 n^2}{\partial r^2}\right)\right.$$

$$\left. - \left(\frac{1}{2\left(n^2 - Q^2\right)}\frac{\partial n^2}{\partial r} - \frac{1}{r}\right)\left(\frac{\delta r}{2}\frac{\partial n^2}{\partial r} - Q\delta Q\right)\right] \qquad (44)$$

and

$$\frac{d(\delta r)}{d\theta} = \frac{r}{\sqrt{n^2 - Q^2}}\left[\frac{\delta r Q}{r} + \delta Q\right.$$

$$\left. - \frac{Q}{n^2 - Q^2}\left(\frac{\delta r}{2}\frac{\partial n^2}{\partial r} - Q\delta Q\right)\right]. \qquad (45)$$

Since power flows within the confines of the rays emanating from a source, the above deviation equations are most useful in calculating the change in power density along a ray path. Lateral to the ray plane, the deviation can be estimated by assuming the rays to follow great-circle paths through the origin of the main ray. However, this approximation is only valid when the lateral variations in the ionosphere are weak.

5. Analytic Techniques

Consider the variational principle for a two-dimensional ray path in Cartesian coordinates (x, y). For the case where the refractive index depends only on the y coordinate, the variational principle becomes

$$\delta \int_T^R n(y)\left[\left(\frac{dy}{dx}\right)^2 + 1\right]^{1/2} dx = 0, \qquad (46)$$

from which a standard result of variational calculus yields a first integral of the Euler-Lagrange equations

$$n(y)\Big/\left[\left(\frac{dy}{dx}\right)^2 + 1\right]^{1/2} = C, \qquad (47)$$

where C is a constant of integration. This is Snell's law for a horizontally stratified ionosphere, and is a single first-order ordinary differential equation for the ray path. In the case of radial coordinates with the refractive index depending on the radial coordinate r alone, there is a first integral of the form $n(r)r^2(d\theta/ds) = C$. This is Bouger's law, which is the generalization of Snell's law to a spherically stratified ionosphere (see [11]). In [12], it was shown that important quantities such as ground range could be derived analytically in the case of an ionospheric layer with refractive index $n = \sqrt{\alpha + \beta/r + \gamma/r^2}$ for $r_b < r < r_m + y_m$, and $n = 0$ elsewhere. Parameters α, β, and γ are given by $\alpha = 1 - (\omega_c/\omega)^2 + (r_b\omega_c/y_m\omega)^2$, $\beta = 2r_b^2 r_m \omega_c/\omega$, and $\gamma = (r_m r_b \omega_c/y_m\omega)^2$, where y_m is the thickness of the layer, r_b is the radius of the base of the layer, r_m is the radius of minimum refractive index, and ω_c is the plasma frequency at this height. Such an ionospheric layer is known as a quasi-parabolic layer. It was shown in [12] that a ray launched at the surface of the Earth with an initial elevation ϕ_0 will land at a distance

$$D = 2r_E\left[(\phi - \phi_0)\right.$$

$$\left. - \frac{r_E\cos\phi_0}{2\sqrt{\gamma}}\ln\frac{\beta^2 - 4\alpha\gamma}{4\gamma\left(\sin\phi + \frac{1}{r_b}\sqrt{\gamma} + \frac{1}{2\sqrt{\gamma}}\beta\right)^2}\right], \qquad (48)$$

where $\cos\phi = r_E/r_b\cos\phi_0$, and r_E is the radius of the Earth. Such results arguably provide the most efficient form of ray tracing.

Snell's law can be further generalized [13] to the following result. If, in two dimensions, the refractive index has the form

$$n(x,y) = R\left(\Im\{g(z)\}\right)\left|g'(z)\right|, \tag{49}$$

where $z = x + jy$, then the ray trajectories will satisfy

$$\frac{R\left(\Im\{g(z)\}\right)}{\left|g'(z)\right|}\frac{d\Re\{g(z)\}}{ds} = C\,. \tag{50}$$

When $g(z) = j\ln z$, we obtain Bouger's law on noting that $\ln z = \log r + j\theta$. Consider the conformal transformation $Z = g(z)$, where $Z = X + jY$. Then, Equation (50) will take the form

$$R(Y)(dX/dS) = C \tag{51}$$

in the new (X,Y) coordinates, with $dS = \left|g'(z)\right| ds$ being the distance element in these new coordinates. Equation (51) can be integrated [9] to yield

$$X + C_0 = \int \frac{C}{\sqrt{R^2(Y) - C^2}} dY\,. \tag{52}$$

We can effectively study the propagation through an ionosphere defined by Equation (49) by studying propagation through a horizontally stratified ionosphere. The quasi-parabolic ionosphere is obtained by introducing $R(Y) = \sqrt{\gamma + \beta\exp(Y) + \alpha\exp(2Y)}$. The integration of Equation (2) then yields

$$X + C_0 =$$

$$\frac{-C}{\sqrt{\gamma - C^2}}\left\{\ln\left[2\sqrt{\gamma - C^2}\sqrt{\gamma - C^2 + \beta\exp(Y) + \alpha\exp(2Y)}\right.\right.$$

$$\left.\left. + \beta\exp(Y) + 2\left(\gamma - C^2\right)\right] - Y\right\}. \tag{53}$$

To obtain the ray path in polar coordinates, we choose $g(z) = j\ln z$, from which we note that $X = -\theta$ and $Y = \ln r$. However, to obtain the results described in [11] we need to add the straight-line sections of the ray that connect the Earth's surface to the base of the ionosphere.

If we choose $g(z) = j\ln(z - z_0)$, instead of $g(z) = i\ln z$, we then obtain the eccentric ionospheres discussed in [13]. Another interesting example arises when $g(z) = \left[\cos(\alpha) + j\sin(\alpha)\right]z$. In this case, the two-dimensional horizontally stratified ionosphere with $\mu(y) = R(y)$ is given a tilt through angle α.

6. Conclusion

We have derived the Haselgrove ray-tracing equations directly from Maxwell's equations, and we have considered some methods for their numerical solution. Even with the fast computers of today, there are still applications for which the solution of the complete Haselgrove equations is still not fast enough. We have therefore looked at some simplifications that could deliver the required speed. In particular, we have looked at ray equations that ignore the background magnetic field and assume propagation to be in a plane. This reduces the ordinary differential equation system to two equations, rather than the six equations of the full Haselgrove formulation. For further speed increases, we have looked at generalizations of Snell's law and the analytic solutions that can be obtained from such generalizations.

7. References

1. J. Haselgrove, "Ray Theory and a New Method for Ray Tracing," *Proc. Phys. Soc.*, The Physics of the Ionosphere, 1954, pp. 355-64.
2. C. B. Haselgrove and J. Haselgrove, "Twisted Ray Paths in the Ionosphere," *Proc. Phys. Soc.*, **75**, 1960, pp. 357-63.
3. J. Haselgrove, "The Hamilton Ray Path Equations," *J. Atmos. Terr. Phys.*, **2**, 1963, pp. 397-9.
4. C. J. Coleman, "A General Purpose Ionospheric Ray Tracing Procedure," Defence Science and Technology Organisation Australia, Technical Report SRL-0131-TR, 1993.
5. D. S. Jones, *Methods in Electromagnetic Wave Propagation*, Oxford, Clarendon Press, 1994.
6. K. G. Budden, *The Propagation of Radio Waves*, New York, Cambridge University Press, 1985.
7. J. R. Smith, *Introduction to the Theory of Partial Differential Equations*, London, Van Nostrand, 1967.
8. J. F. Mathews, *Numerical Methods for Computer Science, Engineering and Mathematics*, New York, Prentice Hall, 1987.
9. C. J. Coleman, "A Ray Tracing Formulation and its Application to Some Problems in OTHR," *Radio Science*, **33**, 1998, pp. 1187-1197.
10. J. A. Bennett, J. Chen, and P. L. Dyson, "Analytic Ray Tracing for the Study of HF Magneto-Ionic Radio Propagation in the Ionosphere," *Applied Computational Electromagnetics Society Journal*, **6**, 1991, pp. 192-210.
11. J. M. Kelso, "Ray Tracing in the Ionosphere," *Radio Science*, **3**, 1968, pp. 1-12.
12. T. A. Croft and H. Hoogasian, "Exact Ray Calculations in a Quasi-Parabolic Ionosphere with no Magnetic Field," *Radio Science*, **3**, 1968, pp. 69-74.
13. C. J. Coleman, "On the Generalization of Snell's Law," *Radio Science*, **39**, 2004.
14. K. Folkestad, "Exact Ray Computations in a Tilted Ionosphere with No Magnetic Field," *Radio Science*, **3**, 1968, pp. 81-84.

Ray Tracing of Magnetohydrodynamic Waves in Geospace

A.D.M. Walker

Abstract

The method of ray tracing is reviewed for MHD waves in stationary media and for media in the steady state. The method is generalized to allow for media that change arbitrarily but slowly in space and time. This requires an additional equation representing the change in frequency along the ray as a result of Doppler shifts. Explicit expressions for ray tracing of transverse Alfvén waves and magnetosonic waves are presented. Some illustrative examples are presented.

1. Introduction

For more than fifty years, since its introduction by Haselgrove [1], ray tracing has been a powerful tool for studying radio propagation in the magnetosphere [2]. This has not been the case for lower-frequency magnetohydrodynamic waves, the chief reason being that often the wavelength is comparable to the size of the magnetosphere, and in no sense can the medium be regarded as slowly varying – a *sine qua non* for the validity of the ray-tracing approximation. For shorter-period waves in the Pc3 range (periods between 10 and 45 seconds), ray tracing may be valid in the magnetosheath and solar wind, while in the solar wind it may be valid for much longer periods. Another difficulty arises in these regions: the medium is moving. Ray tracing in a moving medium has been discussed for sound waves by Lighthill [3], for example, and his discussion was generalized for MHD waves by Walker [4, pp. 429-430.]. The treatment in this paper follows these treatments for steady-state motions, but goes beyond them in allowing for non-steady flow and the resulting Doppler shifts. We do not apply the methods to specific problems of ULF waves in geospace, which will be the subject of a later paper, but instead provide some illustrative numerical examples.

As an example of the issues that are likely to arise in such geospace applications, consider a transverse Alfvén wave. It is well known that the group velocity of such a wave in a stationary medium is directed precisely along the magnetic field, and, since the medium is nondispersive, the group velocity is equal to the Alfvén velocity. The problem of ray tracing in such a case is trivial: the ray paths coincide with the magnetic field lines. However, in a moving medium a wave packet travels with a velocity that is the resultant of the Alfvén velocity and the velocity of the medium. The solar-wind velocity is generally much larger than the Alfvén velocity, so that the ray paths are more nearly radial than along the magnetic-field direction. The case of the magnetosonic waves is even more complicated.

In this paper, we first discuss the geometrical optics of MHD waves in uniform media. We then review the derivation of the ray-tracing equations in moving media in the steady state. Finally, we show how these equations can be generalized to include arbitrary flows, provided changes in the background flow with time are slow compared to the period of the wave. We present explicit equations to allow the numerical computations of MHD rays, and perform some computations to illustrate the techniques.

2. Geometrical Optics of MHD Waves in Uniform Media

2.1 Stationary Media

In a uniform stationary compressible plasma, at frequencies much less than the lowest ion gyrofrequency, there are three characteristic magnetohydrodynamic waves that can be propagated. These can be identified as the transverse Alfvén wave, with the dispersion relation

A. D. M. Walker is with the School of Physics, University of KwaZulu-Natal, Durban 4000, South Africa; e-mail: walker@ukzn.ac.za.

$$\omega = \pm \mathbf{k} \bullet \mathbf{V}_A \qquad (1)$$

or

$$\omega^2 = k^2 V_A^2 \cos^2\theta\,, \qquad (2)$$

and the fast and slow magnetosonic waves, with the dispersion relation

$$\omega^4 - k^2\left(V_A^2 + V_S^2\right) + k^2 V_S^2 \left(\mathbf{k} \bullet \mathbf{V}_A\right)^2 = 0 \qquad (3)$$

or

$$\omega^4 - k^2\left(V_A^2 + V_S^2\right)\omega^2 + k^4 V_A^2 V_S^2 \cos^2\theta = 0\,, \qquad (4)$$

where V_A is the Alfvén speed, V_S is the sound speed, and θ is the angle between the direction of the magnetic field and the wave vector, **k** [4, Equations (7.5) and (7.6)].

These dispersion relations have a special property. They show that the phase velocity, $V_P \equiv \omega/k$, is a function of θ but not of frequency, ω. The propagation of the waves is therefore anisotropic (the phase speed depends on direction) but nondispersive (the phase speed does not depend on frequency). Anisotropic waves have a group velocity (and hence a direction of energy propagation) that in general is not parallel to the wave vector, which is normal to the wavefronts.

As described, for example, by Walker [2], propagation of waves in uniform media can be understood in terms of the properties of three characteristic surfaces: (i) The *wave-vector surface*, given by the magnitude of the wave vector as a function of its direction as represented by the polar angles, θ and ϕ; (ii) the *group-velocity surface*, given by the magnitude of the group velocity as a function of its direction; (iii) the *ray surface*, given by the magnitude of the ray velocity as a function of the direction of energy propagation. The ray velocity is defined as having a magnitude equal to the speed at which wavefronts move along the direction of the group velocity. In the special case of MHD waves, because they are nondispersive, the component of the group velocity in the direction of the wave normal is equal to the phase velocity, and the ray- and group-velocity surfaces are identical.

In radio propagation theory, it is useful to normalize the **k** vector in the form of a refractive index vector, $\mathbf{n} = c\mathbf{k}/\omega$. In MHD, the phase speed, ω/k, is several orders of magnitude less than the speed of light, *c*, so that this leads to refractive indices of the order of 10^3 or 10^4. It is better to express the normalization in terms of a different characteristic speed. The Alfvén speed,

$$V_A \equiv B/\sqrt{\mu_0 \rho}\,, \qquad (5)$$

is suitable. We define a "refractive-index" vector, or normalized wave vector,

$$\mathbf{n} \equiv V_A \mathbf{k}/\omega\,. \qquad (6)$$

The dispersion relations of Equations (2) and (4) then become

$$n\cos\theta = \pm 1\,, \qquad (7)$$

$$n^4 U_S^2 \cos^2\theta - n^2\left(1 + U_S^2\right) + 1 = 0\,,$$

where $U_S \equiv V_S/V_A$. The refractive index is a function of the polar angle, θ, and independent of the azimuthal angle, ϕ. The refractive-index surface in a stationary medium is therefore a surface of rotation about the magnetic-field direction. In particular, the surface for the Alfvén wave, represented by the first equation, is simply a pair of planes where the component of the refractive-index vector parallel to the field is unity. The refractive-index surfaces for the three characteristic waves are shown in Figure 1. The fast wave has the smallest refractive index of the three waves for any direction of propagation. It is a closed surface with an ovid shape. The slow wave has the largest refractive index. Propagation is not possible for directions approximately perpendicular to the magnetic-field direction, so there are two separate surfaces of rotation that are asymptotic to the pair of cones defined by

$$\cos\theta_{res} = \pm U_S^{-1} \equiv V_A/V_S\,. \qquad (8)$$

The two planes representing the transverse Alfvén wave lie between these two surfaces, with $n_{||} = 1$.

The group velocity is given by Equation (2),

$$\mathbf{V}_G = \nabla_k \omega\,, \qquad (9)$$

and thus it is normal to the refractive-index surfaces, which are surfaces of constant ω in *k* space. We immediately see that for the transverse Alfvén wave, the group velocity, and hence the direction of energy propagation, is exactly along the magnetic-field direction. For the fast wave, although it is not parallel to the wave normal, it can make any angle with the magnetic field, so that energy can be propagated in any direction through the medium of the fast wave. For the slow wave, the direction of energy propagation generally makes a small angle with the magnetic field, so that energy is always propagated approximately in the direction of the field and cannot be propagated transverse to it.

The ray surface for the transverse Alfvén wave is degenerate. No matter what the direction of the wave normal, the ray velocity is exactly equal in magnitude to the Alfvén speed, and is directed along the magnetic field. The ray- and group-velocity surfaces are therefore a pair of

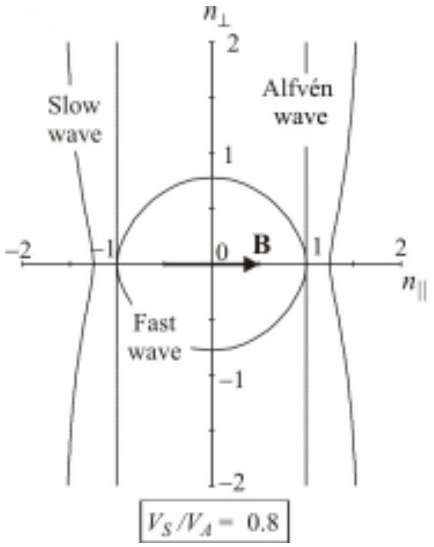

Figure 1a. The refractive-index surface for magnetohydrodynamic waves when the sound speed is less than the Alfvén speed.

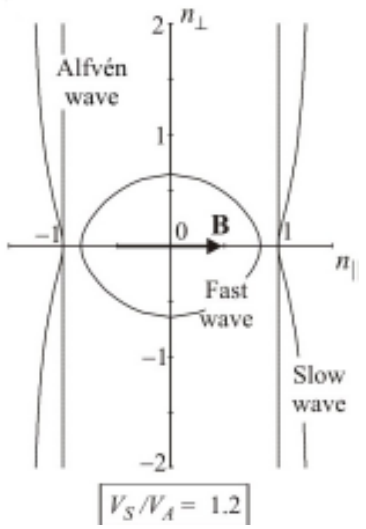

Figure 1b. The refractive-index surface for magnetohydrodynamic waves when the sound speed is greater than the Alfvén speed.

points located a distance $\pm V_A$ from the origin in the direction of the magnetic field.

Examples of ray surfaces for the magnetosonic waves are shown in Figure 2. The surface for the fast wave in the upper panel is an ovoid, showing the magnitude of the group velocity (equivalent in this nondispersive case to the ray velocity) as a function of direction. The surface for the slow wave is more complicated. It can be understood by considering Figure 1a, which shows the corresponding refractive-index surface. For slow-wave propagation with the wave normal parallel to the field in the positive direction, the ray direction normal to the surface is also parallel to the field. As we increase the angle between the wave normal and the magnetic field anticlockwise, the angle between the ray direction and magnetic field increases clockwise, until it reaches a maximum value where there is a point of inflection in the refractive-index surface. At this point, there is a cusp in the ray surface. As the wave-normal angle increases further, the ray direction again decreases. Both ray- and refractive-index surfaces are surfaces of rotation about the magnetic-field direction.

In MHD propagation, a point source radiates waves having the shape of the ray surface. The ray surface also represents the shape of a pulse spreading out from the source, since the medium is nondispersive and wavefronts and signal fronts have the same shape. The ray surface is the correct surface to use in Huygens' construction.

2.2. Moving Media

In uniform moving media, the situation is complicated. Suppose the plasma moves in the reference frame of the observer with velocity **V**. Let the frequency of the wave in the plasma rest frame be ω_0. The frequency in the observer's frame is then Doppler shifted so that

$$\omega_0 = \omega - \mathbf{k} \bullet \mathbf{V} \cdot \qquad (10)$$

The dispersion relations are then those for a stationary medium, Equations (2) and (4), with ω replaced by $\omega - \mathbf{k} \bullet \mathbf{V}$

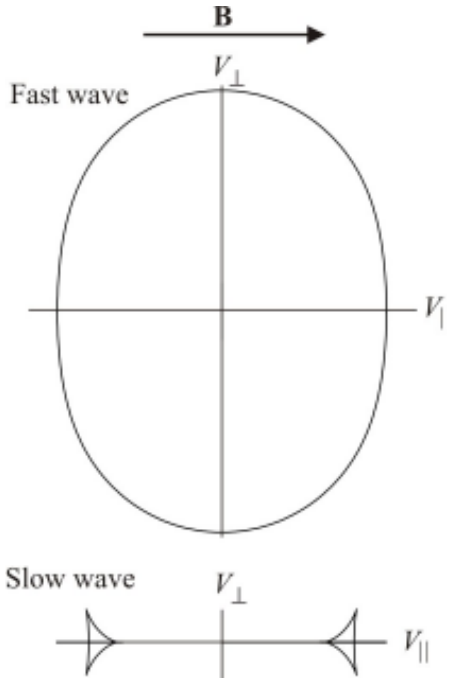

Figure 2. Ray surfaces for the fast and slow waves.

In the moving medium, there is no longer a single axis of rotational symmetry. There are two characteristic axes parallel to the magnetic field and the flow velocity. The refractive-index surfaces are now reflection symmetric across the plane defined by **B** and **V**.

The ray is still normal to the refractive-index surface. The group velocity is

$$\mathbf{V}_G = \nabla_k \omega$$

$$= \nabla_k \left(\omega_0 + \mathbf{k} \bullet \mathbf{V}\right) \qquad (11)$$

$$= \nabla_k \omega_0 + \mathbf{V}$$

$$= \mathbf{V}_{G,0} + \mathbf{V},$$

giving the fairly obvious result that the group velocity in the moving medium is the resultant of the group velocity in the rest frame and the velocity of the medium. In the particular case of the transverse Alfvén wave,

$$\mathbf{V}_G = \mathbf{V}_A + \mathbf{V}, \qquad (12)$$

and the group velocity in the moving medium is not parallel to **B**. For example, in observations of the solar wind by satellites at the solar libration point, such as WIND or ACE, since $\mathbf{V} >> \mathbf{V}_A$, transverse Alfvén waves are propagated almost radially from the sun, rather than parallel to the magnetic field.

3. Theory of Ray Tracing

3.1 Stationary Media

Ray tracing applies when a medium is slowly varying. By this it is meant that the medium does not change appreciably within a wavelength. The ray-tracing equations in a stationary medium – as shown, for example, by Walker [2] – are given by

$$\frac{d\mathbf{r}}{dt} = \nabla_k \omega, \qquad (13)$$

$$\frac{d\mathbf{k}}{dt} = -\nabla \omega,$$

or in Cartesian tensor notation,

$$\frac{dx_i}{dt} = \frac{\partial \omega}{\partial k_i}, \qquad (14)$$

$$\frac{dk_i}{dt} = -\frac{\partial \omega}{\partial x_i},$$

where the subscripts take the values 1, 2, 3, corresponding to the three Cartesian coordinate directions, and summation on the repeated subscript is understood. Precise evaluations of the errors depend on the specific circumstances of the problem. Examples of such error calculations were given by Walker [4, Section 15.4.2] for a particular case, for example. Such calculations show that even when the medium changes by as much as 10% per wavelength, the error in phase is less than one radian in 200π wavelengths. The accuracy of the ray tracing depends on the precision with which the phase is known, and in most ray-tracing situations, the medium varies far more slowly than this.

Let us first consider the trivial case of transverse Alfvén waves, and show that the ray-tracing equations do, in fact, give rays following the field lines with the Alfvén speed. From Equation (1), with the choice of the positive sign, we get

$$\frac{d\mathbf{r}}{dt} = \nabla_k \left(\mathbf{k} \bullet \mathbf{V}_A\right) = \mathbf{V}_A, \qquad (15)$$

$$\frac{d\mathbf{k}}{dt} = -\nabla \left(\mathbf{k} \bullet \mathbf{V}_A\right) \cdot$$

As expected, the first of these shows that the wave packet moves parallel to $\mathbf{V}_A$, and hence to **B**. (The choice of the other sign would have given a wave packet moving anti-parallel to **B**). The second equation shows how the wavenumber, and hence the wavelength, change along the ray but are independent of the first equation. Thus, rather than two simultaneous vector differential equations for the ray path, we can evaluate the change of ray path and the change of wavelength independently. For the magnetosonic waves, we evaluate the expressions on the right-hand side of Equation (15) explicitly using Equation (4). The result, in Cartesian tensor notation, is

$$\frac{dx_i}{dt} = \frac{\left[k_i \left(V_A^2 + V_S^2\right)\omega^2 - k_i \left(k_j V_{A,j}\right)^2 V_S^2\right]}{\omega\left[2\omega^2 - k^2\left(V_A^2 + V_S^2\right)\right]}$$

$$- \frac{\left(k_j V_{A,j}\right) k^2 V_S^2 V_{A,i}}{\omega\left[2\omega^2 - k^2\left(V_A^2 + V_S^2\right)\right]}, \qquad (16)$$

$$\frac{dk_i}{dt} = -k^2 \frac{\omega^2 \left[V_{A,j} \frac{\partial V_{A,j}}{\partial x_i} + V_S \frac{\partial V_S}{\partial x_i} \right]}{\omega \left[2\omega^2 - k^2 \left(V_A^2 + V_S^2 \right) \right]}$$

$$- \frac{k_j k_k V_{A,j} V_S^2 \frac{\partial V_{A,k}}{\partial x_i} + \left(k_j V_{A,j} \right)^2 V_S \frac{\partial V_S}{\partial x_i}}{\omega \left[2\omega^2 - k^2 \left(V_A^2 + V_S^2 \right) \right]} .$$

The model provides expressions for the variation of $\mathbf{V}_A$ and V_S with position. These six simultaneous first-order equations can be integrated numerically by a suitable step-by-step process, such as the Runge-Kutta method, to provide the path of the ray.

3.2 Steady-State Flow

In steady-state flow, the velocity of the medium is a function of position, but it is independent of time, so that at any point in space the velocity remains constant. In this case, we can use Equation (10) and write the ray-tracing equations in the form

$$\frac{d\mathbf{r}}{dt} = \nabla_k \omega = \nabla_k \left(\omega_0 + \mathbf{k} \bullet \mathbf{V} \right) = \nabla_k \omega_0 + \mathbf{V} , \quad (17)$$

$$\frac{d\mathbf{k}}{dt} = -\nabla \omega = -\nabla \left(\omega_0 + \mathbf{k} \bullet \mathbf{V} \right),$$

or in subscript notation,

$$\frac{dx_i}{dt} = \frac{\partial \omega_0}{\partial k_i} + V_i , \quad (18)$$

$$\frac{dk_i}{dt} = -\frac{\partial \omega_0}{\partial x_i} - k_j \frac{\partial V_j}{\partial x_i} .$$

This shows that – as for a uniform medium – the wave packet moves with a group velocity that is the resultant of the group velocity in the local rest frame of the medium and the velocity of the medium at that point. The ray-tracing equations are then given by Equation (16), with ω replaced by ω_0 and the addition of the terms V_i and $-k_j \, \partial V_j / \partial x_i$, which can be evaluated from the model.

This set of equations allows us to follow the path of a wave packet. The frequency in the frame in which the motion is steady state is ω and is constant. (This is justified in the more general treatment of the next section). The frequency in the local rest frame is Doppler shifted, and varies from point to point. This has consequences for the wave energy: as the wave progresses, it exchanges energy with the kinetic energy of the background flow.

3.3 The General Case

The most general case that can be treated by ray-tracing methods is one in which the properties of the medium are functions of both position and time, and no frame of reference can be found in which the medium has a steady-state flow. The restriction applied is that the definition of a slowly varying medium must be extended. The medium is assumed to vary slowly in space such that length scales are long compared with the wavelength and, in addition, the time scale of variation is long compared with the wave's period. The complication now introduced is that the frequency can no longer be taken as constant. There will now be Doppler shifts in the frequency as the wave progresses, arising from the change in the properties of the medium as a function of time.

The dispersion relation can now be written formally as

$$\omega = \omega \left(x_i, k_i, t \right), \quad (19)$$

where there is now an explicit dependence on the time, *t*, as a result of the dependence of the Alfvén velocity and sound speed on time.

We now repeat the derivation of the ray-tracing equations as given by Walker [2], but including the dependence of the dispersion relation on time.

The waves are assumed to vary in space and time as

$$\exp i\Phi \equiv \exp i \left[\int k_i dx_i - \int \omega \left(x_i, k_i, t \right) dt \right] \quad (20)$$

$$= \exp i \int \left[k_i \frac{dx_i}{dt} - \omega \left(x_i, k_i, t \right) \right] dt \cdot$$

The variation of the dispersion relation with position and time is assumed to be slow, so that

$$\frac{\partial \Phi}{\partial t} \approx \omega , \quad (21)$$

$$\frac{\partial \Phi}{\partial x_i} = -k_i \cdot$$

A wave packet can be represented by a Fourier synthesis of plane waves:

$$f(x_i,t)=\int_{-\infty}^{\infty}F(k_1,k_2,k_3)$$

$$\exp\left\{i\int\left[k_i\frac{dx_i}{dt}-\omega(x_i,k_i,t)\right]dt\right\}dk_1dk_2dk_3 \quad (22)$$

The method of stationary phase shows that the only significant contributions occur along paths of constructive interference where the phase is stationary with respect to variations in the components of **k**:

$$\frac{\partial\Phi}{\partial k_i}=0, \quad (23)$$

where Φ is a function of the seven independent variables $x_1, x_2, x_3, k_1, k_2, k_3$, and t. It is understood that the partial differentiation with respect to a component of k_i implies holding not only the other two components of k_i constant, but also the three components of x_i and t constant. If we differentiate the phase in Equation (22) with respect to k_i and equate it to zero, we get

$$\int\left(\frac{dx_i}{dt}-\frac{\partial\omega}{\partial k_i}\right)dt=0. \quad (24)$$

This must hold for integration over an arbitrary time interval so that the integrand must be zero. This gives the first ray-tracing equation:

$$\frac{dx_i}{dt}=\frac{\partial\omega}{\partial k_i}. \quad (25)$$

The second ray-tracing equation is found by using the two Equations (21) to write

$$\frac{\partial\Phi}{\partial t}=\omega\left(x_i,-\frac{\partial\Phi}{\partial x_i},t\right). \quad (26)$$

We differentiate this with respect to x_i, getting

$$\frac{\partial^2\Phi}{\partial x_i\partial t}\equiv-\frac{\partial k_i}{\partial t}$$

$$=\frac{\partial\omega}{\partial x_i}+\frac{\partial\omega}{\partial k_j}\frac{\partial k_i}{\partial x_j} \quad (27)$$

$$=\frac{\partial\omega}{\partial x_i}+V_{G,i}\frac{\partial k_i}{\partial x_j},$$

or

$$\frac{\partial k_i}{\partial t}+V_{G,j}\frac{\partial k_i}{\partial x_j}=-\frac{\partial\omega}{\partial x_i}. \quad (28)$$

The left-hand side of this is just the total derivative dk_i/dt, the time rate of change of k_i following a wave packet moving with the group velocity $V_{G,j}$. We thus get the second ray-tracing equation.

In non-stationary media, the frequency is not constant. We evaluate the rate of change of w following a wave packet:

$$\frac{d}{dt}\left[\omega(x_i,k_i,t)\right]=\frac{\partial\omega}{\partial x_i}\frac{dx_i}{dt}+\frac{\partial\omega}{\partial k_i}\frac{dk_i}{dt}+\frac{\partial\omega}{\partial t}$$

$$=\frac{\partial\omega}{\partial x_i}\frac{\partial\omega}{\partial k_i}-\frac{\partial\omega}{\partial k_i}\frac{\partial\omega}{\partial x_i}+\frac{\partial\omega}{\partial t} \quad (29)$$

$$=\frac{\partial\omega}{\partial t}.$$

In steady-state flow, the dispersion relation is independent of time. Thus, $d\omega/dt=0$, and the frequency remains constant along a ray. If the flow is not steady, then $\partial\omega/\partial t\neq 0$, and this is given by the model. We can then follow the change of frequency along the ray using this third equation. The full set of ray-tracing equations to be integrated is then

$$\frac{dx_i}{dt}=\frac{\partial\omega}{\partial k_i},$$

$$\frac{dk_i}{dt}=-\frac{\partial\omega}{\partial x_i}, \quad (30)$$

$$\frac{d\omega}{dt}=\frac{\partial\omega}{\partial t}.$$

We can use Equation (10) to express this in terms of ω_0, the Doppler-shifted frequency in the local rest frame of the plasma, getting

$$\frac{dx_i}{dt} = \frac{\partial\left(\omega_0 + k_j V_j\right)}{\partial k_i} = \frac{\partial\omega_0}{\partial k_i} + V_i ,$$

$$\frac{dk_i}{dt} = -\frac{\partial\left(\omega_0 + k_j V_j\right)}{\partial x_i} = -\frac{\partial\omega_0}{\partial x_i} - k_j \frac{\partial V_j}{\partial x_i} , \qquad (31)$$

$$\frac{d\omega}{dt} = \frac{\partial\left(\omega_0 + k_j V_j\right)}{\partial t} = \frac{\partial\omega_0}{\partial t} + k_j \frac{\partial V_j}{\partial t} .$$

This is the most general form of the ray-tracing equations. They can be integrated step-by-step to follow the path of a wave packet. In steady-state flows, successive wave packets follow the same path so that the ray for a particular set of starting conditions is fixed in space. If the flow at a point varies with time, then this is not so, and successive wave packets follow different paths. As written here, these equations apply to any type of wave in a moving medium. If ω_0 is given by the dispersion relation of Equation (4), the equations describe magnetosonic waves. In this case, we can write them explicitly as

$$\frac{dx_i}{dt} = \frac{\left[k_i\left(V_A^2 + V_S^2\right)\omega_0^2 - \left(k_j V_{A,j}\right)^2 k_i V_S^2\right]}{\omega_0\left[2\omega_0^2 - k^2\left(V_A^2 + V_S^2\right)\right]}$$

$$- \frac{\left(k_j V_{A,j}\right)k^2 V_S^2 V_{A,i}}{\omega_0\left[2\omega_0^2 - k^2\left(V_A^2 + V_S^2\right)\right]} + V_i ,$$

$$\frac{dk_i}{dt} = -k^2 \frac{\omega_0^2\left[V_{A,j}\dfrac{\partial V_{A,j}}{\partial x_i} + V_S\dfrac{\partial V_S}{\partial x_i}\right]}{\omega_0\left[2\omega_0^2 - k^2\left(V_A^2 + V_S^2\right)\right]}$$

$$- \frac{k_j k_k V_{A,j} V_S^2 \dfrac{\partial V_{A,k}}{\partial x_i} + \left(k_j V_{A,j}\right)^2 V_S \dfrac{\partial V_S}{\partial x_i}}{\omega_0\left[2\omega_0^2 - k^2\left(V_A^2 + V_S^2\right)\right]} - k_j \frac{\partial V_j}{\partial x_i} \qquad (32)$$

$$\frac{d\omega}{dt} = k^2 \frac{\omega_0^2\left[V_{A,j}\dfrac{\partial V_{A,j}}{\partial t} + V_S\dfrac{\partial V_S}{\partial t}\right]}{\omega_0\left[2\omega_0^2 - k^2\left(V_A^2 + V_S^2\right)\right]}$$

$$- \frac{k_j k_k V_{A,j} V_S^2 \dfrac{\partial V_{A,k}}{\partial t} + \left(k_j V_{A,j}\right)^2 V_S \dfrac{\partial V_S}{\partial t}}{\omega_0\left[2\omega_0^2 - k^2\left(V_A^2 + V_S^2\right)\right]} + k_j \frac{\partial V_j}{\partial t}$$

The ray-tracing equations for the transverse Alfvén wave are

$$\frac{dx_i}{dt} = \pm V_{A,i} + V_i ,$$

$$\frac{dk_i}{dt} = \mp k_j \frac{\partial V_{A,j}}{\partial x_i} - k_j \frac{\partial V_j}{\partial x_i} , \qquad (33)$$

$$\frac{d\omega}{dt} = \pm k_j \frac{\partial V_{A,j}}{\partial t} + k_j \frac{\partial V_j}{\partial t} ,$$

where the upper sign corresponds to a wave propagated parallel to, and the lower sign to one propagated anti-parallel to, the magnetic field in the local rest frame.

4. Some Numerical Examples

It is not the intention of this paper to provide detailed numerical solutions to specific magnetospheric problems. This will be the subject of further work. We provide two very simple numerical examples that illustrate the technique, without the complexity of the most general case.

4.1 Isotropic Alfvén Wave in a Shear Layer

Consider a shear layer in a cold plasma (i.e., P and therefore V_S are zero), as illustrated in Figure 3a. The magnetic field is uniform and in the y direction. The density is also uniform, so that the Alfvén speed is constant. The plasma has a velocity in the x direction that varies as a function of z, thus forming a shear layer. This is a steady-state problem, so that w is constant. If we make use of the dispersion relation $\omega_0^2 = k^2 V_A^2$, the ray-tracing equations, Equations (32) for an isotropic Alfvén wave (the fast wave), may be written

$$\frac{dx}{dt} = \frac{k_x V_A^2}{\omega_0} + V(z) ,$$

$$\frac{dy}{dt} = \frac{k_y V_A^2}{\omega_0} ,$$

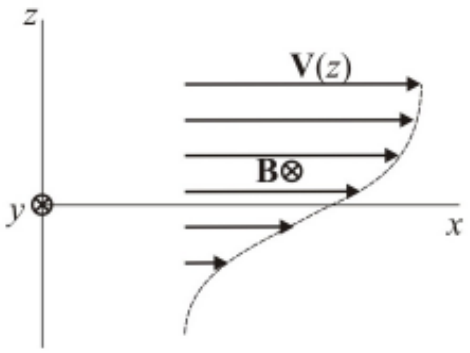

Figure 3a. The geometry of the model of a shear layer.

$$\frac{dz}{dt} = \frac{k_z V_A^2}{\omega_0}, \tag{34}$$

$$\frac{dk_x}{dt} = \frac{dk_y}{dt} = 0,$$

$$\frac{dk_z}{dt} = k_x \frac{dV}{dz}.$$

These show that k_x and k_y are constant. Now suppose that there is a fast wave incident from below the layer in the *x*-*z* plane so that $k_y = 0$. Then *y* is constant, and the wave remains in the *x*-*z* plane. We assume that *V*, the velocity in the shear layer, varies with *z* as

$$V(z) = \tfrac{1}{2} V_0 \left(\tanh \frac{z}{w} + 1 \right). \tag{35}$$

Then,

$$\omega_0 = \omega - k_x V(z) = \omega - \tfrac{1}{2} k_x V_0 \left(\tanh \frac{z}{w} + 1 \right), \tag{36}$$

and

$$\frac{dV}{dz} = \frac{V_0}{2w} \operatorname{sech}^2 \frac{z}{w}. \tag{37}$$

The ray-tracing equations can then be written

$$\frac{dx}{dt} = \frac{k_x V_A^2}{\omega - k_x V(z)} + V(z),$$

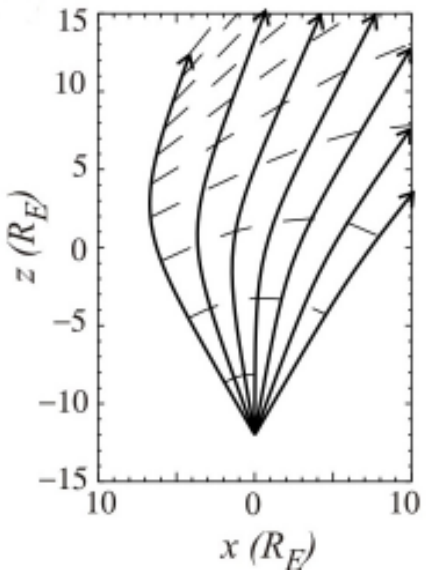

Figure 3b. Rays with various angles of incidence traced through a shear layer. The dashed lines represent lines of constant phase. $V_A = 0.031\ \mathrm{R_E/s}$, $\omega = 0.31$ radians/s.

$$\frac{dz}{dt} = \frac{k_z V_A^2}{\omega - k_x V(z)}, \tag{38}$$

$$\frac{dk_z}{dt} = \frac{k_x V_0}{2w} \operatorname{sech}^2 \frac{z}{w},$$

where $V(z)$ is given by Equation (35). For the ray tracing to be valid, the medium must be slowly varying, which means that *w*, the half-width of the layer, must be large

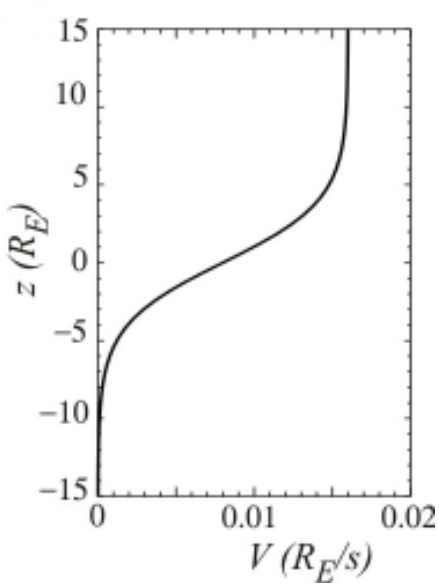

Figure 3c. The magnitude of the shear velocity in the x direction as a function of z given by $V(z) = 0.008\left\{\tanh\left(\tfrac{1}{4}z\right)+1\right\}$, *where z is in* $\mathrm{R_E}$ *and V is in* $\mathrm{R_E/s}$.

compared with the wavelength, i.e.,

$$kw \equiv \frac{\omega}{V_A} w >> 1 \cdot \qquad (39)$$

In Figure 3b, we show the results of a numerical integration of these equations. We used a Runge-Kutta method to find the rays originating at a point source well below the shear layer, where the velocity of the medium is essentially zero. We traced seven rays, initially making angles with the z axis from $-30°$ to $+30°$ at $10°$ intervals. Figure 3c shows the actual velocity profile that was used as a function of z. If the velocity had been zero everywhere, then the rays would have been straight lines, and the wavefronts would have been spherical. The dashed lines in Figure 3b are contours of equal times of propagation of the wave packet from the source. As discussed above, the ray velocity and the group velocity are the same for MHD waves and, therefore, these are surfaces of constant phase representing wavefronts or wave crests. The diagram shows graphically how the wave is swept along by the increasing velocity shear as it is propagated.

4.2 Magnetic Field Increasing in Time

We now consider a different problem, in which the system is not in a steady state, and some of the parameters vary in time. Consider a uniform region of plasma, containing a sheet of current flowing in the y direction, and situated on the plane $z = 0$. Associated with the current sheet is a uniform magnetic field, which for $z > 0$ is in the x direction, and for $z < 0$ is in the $-x$ direction. This can be regarded as a crude model of the magnetotail near its axis. We tackle the problem of a transverse Alfvén wave propagated along the x axis while this system changed in time.

Suppose that the surface current density in the current sheet increases slowly and quasi-statically, by which we mean that any acceleration is so small that the plasma remains in quasi-equilibrium as conditions change. The magnetic field increases uniformly, and the EMF arising from the changing magnetic flux provides an electric field associated with a plasma drift inwards to the current sheet, at a velocity consistent with the frozen-in-field-line theorem. This crudely simulates the changes that occur when an increase in reconnection rate at the sub-solar point leads to an increase of flux in the tail lobes.

Consider a rectangular block of plasma as shown in Figure 4, bounded between two planes located at the current sheet and at a height z above it. The magnetic field is a function only of t. At time t, the magnetic flux through the plane $CDD'C'$ is Bwz. Since B is a function of time, there is an EMF around the loop $CDD'C'$. From the symmetry of the problem, this is associated with an electric field that is in the y direction. From Faraday's law, the EMF around the loop is minus the time rate of change of B. Thus,

$$Ew = wz \frac{dB}{dt} \qquad (40a)$$

or

$$E(z,t) = z \frac{dB(t)}{dt} \cdot \qquad (40b)$$

This electric field leads to an $\mathbf{E} \times \mathbf{B}$ drift of plasma towards the current sheet in the $-z$ direction, with magnitude

$$V(z,t) = \frac{E(z,t))}{B(t)} = \frac{z}{B(t)} \frac{dB(t)}{dt} \cdot \qquad (41)$$

This drift is consistent with the frozen-in field theorem. The density of plasma in the block therefore increases uniformly. The mass of plasma flowing downwards through the top of the block during time dt is

$$dm = lw\rho(t)V(z,t)dt$$

$$= \frac{\rho lwz}{B} \frac{dB}{dt} dt \qquad (42)$$

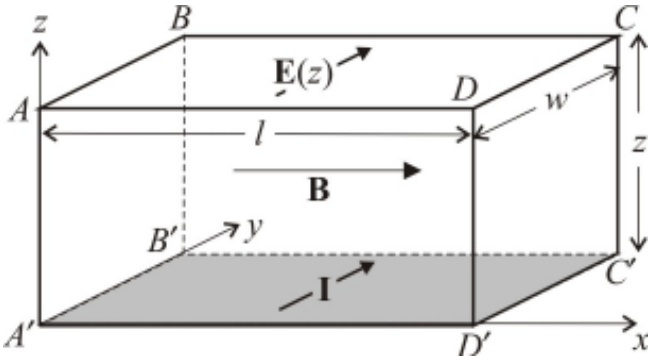

Figure 4. A region of increasing magnetic field: see the text for an explanation.

$$= \frac{m}{B}\frac{dB}{dt}\,.$$

Thus,

$$\frac{1}{m}\frac{dm}{dt} \equiv \frac{1}{\rho}\frac{d\rho}{dt} = \frac{1}{B}\frac{dB}{dt}\,. \qquad (43)$$

Now, consider a transverse Alfvén wave with initial frequency ω_0, propagated in the x direction with $k_x = k$ and $k_y = k_z = 0$. The ray-tracing equations for such a wave are given by Equation (33). The equations for the rate of change of the components of the wave vector **k** show that since k_y and k_z are initially zero, they remain zero, and since V_A is independent of x, k_x remains constant and equal to k. Thus, the wave-normal direction and the wavelength are constant. For the geometry we are considering, the ray-tracing equations may therefore be written

$$\frac{dx}{dt} = V_A\,,$$

$$\frac{dz}{dt} = -V\,, \qquad (44)$$

$$\frac{d\omega}{dt} = k_x\frac{dV_A}{dt}\,.$$

Two of these equations can be integrated directly. From Equation (41), we see that

$$\frac{1}{z}\frac{dz}{dt} = \frac{1}{B}\frac{dB}{dt}\,, \qquad (45)$$

so that z/B is constant, and a wave packet starting at z_0 has its z coordinate changing as

$$z(t) = B_0 z_0 / B(t)\,. \qquad (46)$$

The last of the ray-tracing Equations (44) for w may be directly integrated to get

$$\omega(t) = kV_A(t)\,. \qquad (47)$$

This shows that since the wavelength remains constant and the Alfvén speed varies with time, the dispersion relation requires the frequency to have the same time dependence as V_A. The first of Equations (44) shows that

$$x = x_0 + \int_{t_0}^{t} V_A(t)\,dt\,. \qquad (48)$$

To take a specific example, suppose the magnetic field changes with time such that

$$B(t) = B_0 + \tfrac{1}{2}\left(B_f - B_0\right)\left[\tanh\left(\frac{t}{T}\right)+1\right] \qquad (49)$$

$$= B_0\left\{1+\tfrac{1}{2}f\left[\tanh\left(\frac{t}{T}\right)+1\right]\right\},$$

where B_0 and B_f are the initial and final values of B, and f is the fractional change in B. From Equation (43), we see that the plasma density also changes as plasma is transported inwards by the $\mathbf{E}\times\mathbf{B}$ drift, so that

$$\rho(t) = \rho_0\left\{1+\tfrac{1}{2}f\left[\tanh\left(\frac{t}{T}\right)+1\right]\right\}. \qquad (50)$$

The Alfvén speed, Equation (5), therefore depends on time as

$$V_A = V_{A,0}\sqrt{1+\tfrac{1}{2}f\left[\tanh\left(\frac{t}{T}\right)+1\right]}\,. \qquad (51)$$

The downward drift velocity from Equation (41) is then

$$V(z,t) = \frac{zf\,\mathrm{sech}^2\left(t/T\right)}{2T\left\{1+\tfrac{1}{2}f\left[\tanh\left(t/T\right)+1\right]\right\}}\,. \qquad (52)$$

From Equations (46) and (48), the ray path is then given by

$$x(t) = x_0 + \int_{-\infty}^{t} V_{A,0}\sqrt{1+\tfrac{1}{2}f\left[\tanh\left(\frac{t}{T}\right)+1\right]}\,dt \qquad (53)$$

$$z(t) = \frac{z_0}{\left\{1+\tfrac{1}{2}f\left[\tanh\left(\frac{t}{T}\right)+1\right]\right\}},$$

while the frequency varies as

$$\omega(t) = \omega_0\sqrt{1+\tfrac{1}{2}f\left[\tanh\left(\frac{t}{T}\right)+1\right]}\,. \qquad (54)$$

We integrated these equations for this model when the Alfvén speed was initially $0.01R_E\mathrm{s}^{-1}$, with B and ρ

decreasing to half their value in a time such that $T = 1000$ s. This was a much larger change than would be typical of the real magnetotail, but illustrated the physics of the problem. The initial frequency used in the computation was $0.5\ \mathrm{s}^{-1}$. However, since all frequencies have the same group and phase velocity, the ray paths calculated were independent of ω, and the change in frequency was proportional to the initial frequency. Thus, the results applied to any frequency, provided that it was not so low that the slowly-varying criterion was violated.

The results are shown in Figure 5 for a wave packet that started at $22.5R_E$ from the current sheet. Figure 5a shows the ray path, and Figure 5b shows the change in the frequency ω. Ticks are shown along the ray path representing equal intervals of time. As the medium was compressed, the magnitudes of the magnetic field and the density were doubled, and the Alfvén speed increased by a factor of $\sqrt{2}$. Since the wavelength was unchanged, the frequency $f = V_A/\lambda$ was increased by the same factor.

In interpreting this diagram, it is important to understand that there are significant differences from the behavior shown in a stationary or a steady-state medium. If there is no dependence on *t*, successive parts of a wave of a given frequency follow the same path. The ray path is independent of time. When the medium varies in time, parts of the wave passing through a given point follow different paths as the medium changes. To clarify this, consider Figure 6. Imagine that at time $t = 0$, just as the compression is becoming appreciable, we have a continuous transverse Alfvén wave limited in space, forming a thin pencil aligned with the magnetic field and represented by the line *ABCD*. Assume also that the wave normal is in the *x* direction. The points *A*, *B*, *C*, and *D* are equally spaced and separated by an integral number of wavelengths. We consider the location of the signal at a succession of times: t_1, t_2, etc., separated by equal intervals. Each portion of the pencil moves along a ray path like that of Figure 5a. Wave packets that initially followed each other along a single ray path parallel to the field now follow different paths as the medium changes in time. At consecutive equal time intervals, the signal occupies a pencil defined by the change in position of the wave packets at *A*, *B*, *C*, and *D*. The horizontal distance between these points remains constant, and thus the wavelength cannot change. However, the velocity is increasing, as can be seen from the fact that the horizontal distance between successive positions of each point increases. Thus, the frequency must also increase. The net effect of the compression is that the wave, initially confined to a field line and propagated along it, continues to be confined to the same frozen-in field line as it is carried inwards by the compression. As this happens, the velocity and frequency increase.

This is an illuminating example that keeps clear the distinction between temporal and spatial effects. In more complicated cases, spatial and temporal effects are inextricably interconnected.

5. Discussion and Conclusions

Ray tracing has not often been used as a technique for analyzing the behavior of MHD waves in geospace. Frequently, the wavelengths are so large that the medium cannot be regarded as slowly varying, so that full-wave treatments are necessary. In addition, characteristic wave

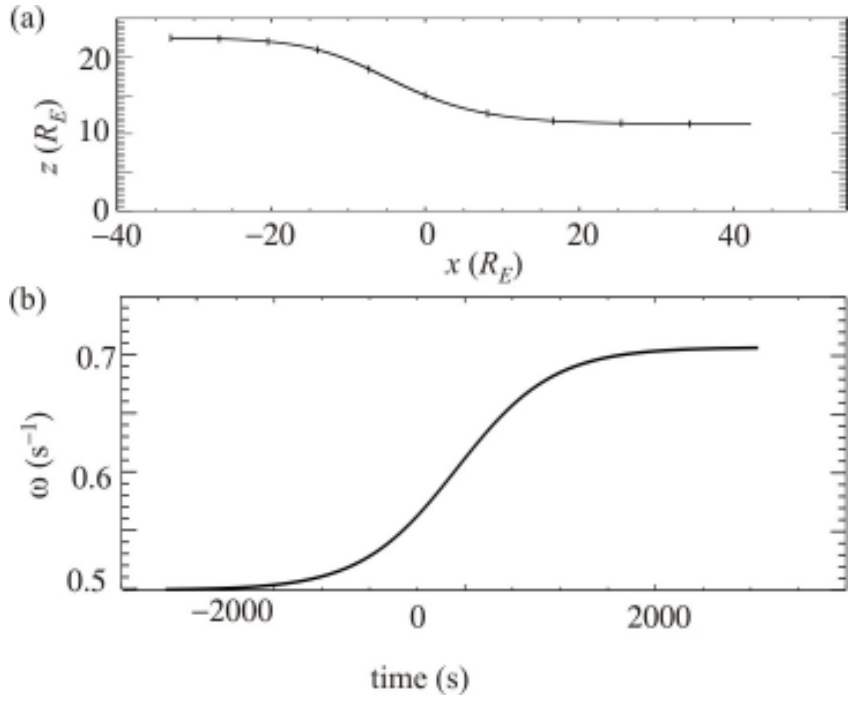

Figure 5. Ray tracing in a model in which the magnetic field is increasing with time: $V_{A,0} = 0.01\ \mathrm{R_E s^{-1}}$, $\omega_0 = 0.5\ \mathrm{s}^{-1}$, $T = 1000$ s, $f = 1.0$. *(a) The ray path of a wave packet. (b) The change in frequency with magnetic field increase as a function of time.*

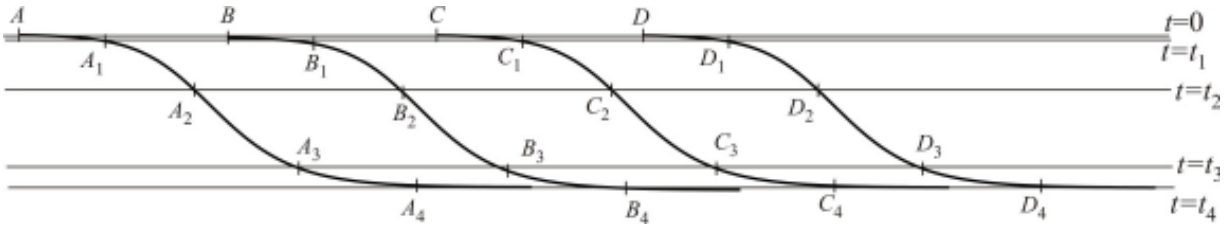

Figure 6. The motion of a ray pencil that is aligned with the magnetic field.

velocities are small, so that the motion of the medium is significant. The velocity of the medium relative to the observer is sufficient to influence the dispersion relation substantially as a result of Doppler shift, whereas at radio frequencies the wave velocities are usually large enough so that the medium can be regarded as being at rest.

In this paper, we have reviewed the ideas of ray tracing for MHD waves in stationary media. We have then generalized them to apply to steady-state motion of the medium, in the same way as has been done for sound waves, for example, by Lighthill [3]. We have also presented equations that allow for the tracing of rays in slowly varying media with motion that is not steady-state flow. This is a new result, showing how the frequency may change along the ray as a result of Doppler shifts.

It is not our intention in this paper to apply the equations to specific magnetospheric problems. However, it is worth considering the circumstances in which they would be useful. One obvious case would be that of Pc3 pulsations originating from the neighborhood of the bow shock. These are propagated from the shock through the magnetosheath, and are eventually transmitted into the ionosphere. Ray-tracing studies of such waves in the past [5] have assumed a stationary medium. The magnetosheath has strong streaming velocities past the magnetopause, but for most purposes could be treated as being in a steady state. The steady-state equations would be appropriate, and the Pc3 wavelengths are short enough for the ray-tracing approximation to be valid. The examples in this paper suggest that this could lead to significant modifications of the results for a stationary medium.

In the solar wind, the flow is supersonic and super-Alfvénic. It may in some regions approximate a steady-state flow, but usually this will not be the case. The appropriate equations would be the general equations. In such a case, the model might often need to be so complicated that computation might not be worth while. Succeeding wave packets would not follow the same path as conditions changed in time. Nevertheless, the equations make explicit the fact that as each wave packet is propagated, its frequency changes according to the third of Equations (31); attempts to correlate spectral peaks observed at different locations may fail because of Doppler shifts.

A topic that we have not discussed is the exchange of energy between background flows and the wave. In the case of atmospheric winds, this can lead to modification of the background flow by the wave, and amplification or attenuation of the wave through the energy exchange. Study of such processes in the solar wind is still to be carried out.

It is hoped that these equations will be applied to the solution of wave-propagation problems in geospace.

6. References

1. J. Haselgrove, "Ray Theory and a New Method for Ray Tracing," in *The Physics of the Ionosphere: Report of the Physical Society Conference on the Physics of the Ionosphere, held at Cavendish Laboratory, Cambridge, September 1954*, London, Physical Society, 1955, pp. 355-364.
2. A. D. M. Walker, "Ray Tracing in the Magnetosphere," *Radio Science Bulletin*, No. 326, September 2008, to appear.
3. J. Lighthill, *Waves in Fluids*, Cambridge, Cambridge University Press, 1978.
4. A. D. M. Walker, *Magnetohydrodynamic Waves in Geospace – The Theory of ULF Waves and their Interaction with Energetic Particles in the Solar-Terrestrial Environment*, Bristol, Institute of Physics Publishing, 2005.
5. X. Zhang, R. H. Comfort, D. L. Gallagher, J. L. Green, Z. E. Musielak, and T. E. Moore, "Magnetospheric Filter Effect for Pc3 Alfvén Mode Waves," *Journal of Geophysical Research – Space Physics*, **100**, 1995, pp 9585-9590.

Practical Applications of Haselgrove's Equations for HF Systems

L.J. Nickisch

Abstract

Jenifer Haselgrove's 1955 paper, entitled "Ray Theory and a New Method for Ray Tracing," has become the classic reference for the application of Hamilton's equations to the problem of ionospheric radiowave propagation. Her formulation provides a computationally tractable means for analyzing radiowave propagation when ionospheric refraction is a significant effect, as is the case in the high-frequency (HF: 3-30 MHz) band, where shortwave communication systems and over-the-horizon (OTH) radars operate. The study of HF propagation effects on communications and radar systems has formed a substantial portion of my career, and so it is with pleasure that I review the more significant aspects of that work to honor Jenifer Haselgrove (now Jenifer Leech) for her very significant contribution, Haselgrove's Equations. I discuss the extension of Haselgrove's Equations for accurate computation of ray focusing and ray homing, and the application of Haselgrove's Equations in HF propagation-channel modeling, simulation of Doppler-spread surface clutter for OTH radar, geolocation of targets for OTH radar, and the mitigation of traveling ionospheric disturbances (TIDs) for OTH radar.

1. Introduction

September 4, 1984, freshly out of graduate school and on the first day of my new job at Mission Research Corporation (MRC) in Santa Barbara, California, I was handed a copy of Budden's classic reference, *Radio Waves in the Ionosphere* [1]. I was told that I was to become an expert on HF propagation. Assured by my boss that nothing useful was expected from me for six months, I was free to delve deeply into the study of HF propagation theory, and devoured "Budden" cover-to-cover. It was there that I learned of Jenifer Haselgrove's derivation of the equations that bear her name, Haselgrove's Equations [2]. As Budden states in his book, referring to the problem of ray tracing in slowly varying magnetoionic media, "Haselgrove has shown how the differential equations for a very general coordinate system can be derived from Fermat's principle of stationary time." I was very familiar with Fermat's principle and variational calculus from my graduate studies, having applied it in my dissertation project to a problem in quantum field theory (not to mention countless problems in my graduate-level classical-mechanics and field-theory courses). Haselgrove's Equations are an implementation of Hamilton's equations, suitable for numerical integration on a computer, which she applied to ionospheric radiowave propagation [3, 4].

I soon acquired the Jones-Stephenson "Raytrace" code [5], the best-known and (at least at the time) most readily available computer program applying Haselgrove's Equations to radiowave propagation in the ionosphere. I came to know this code intimately, and was soon facile in modifying it with subroutines of my own creation.

The Cold War was still strongly in effect in those days, and Mission Research's role as a think tank included estimating the effects of nuclear detonations on the ionosphere, and the subsequent deleterious effects on communication and radar system performance. The first project I was assigned was to compute the effect on the signal strength of HF communication links caused by traveling ionospheric disturbances (TIDs), driven by acoustic gravity waves (AGWs) generated by low-altitude nuclear bursts. This, I was told, could be addressed by tracing a "flux tube" of nearly-separated rays, and computing the ray focusing (or defocusing) by the change in the cross sectional area of the flux tube. However, I soon discovered that in a massive-attack nuclear scenario, the ionosphere could become so convoluted by interfering TIDs that it was impossible to accurately calculate the signal strength in this way. If the flux tube was too large, then the answer was merely an average over the large region intercepted by the flux tube. To get more accuracy, one could make the flux tube smaller, but eventually finite numerical precision

L. J. Nickisch is with NorthWest Research Associates, 301 Webster Street, Monterey, CA 93940 USA; Tel: +1 (831)-582-4905; e-mail: LJ@nwra.com.

would cause meaningless answers as nearly identical ray landing points were differenced in calculating the cross-sectional area. It occurred to me that it should be possible to compute the area of a differential flux tube about a single ray by the integration of differential ray displacements about an undeviated ray. The equations, I knew, could be derived by secondary variations of Haselgrove's Equations. I discuss this formulation in Section 2, where I will further explain how these same equations can be used to efficiently home a ray to a desired landing point (useful, for example, for generating synthetic oblique ionograms, or performing computations on a fixed link).

HF communication systems and radars operating in the nuclear-disturbed ionosphere can be adversely affected by scatter from small-scale ionization structure (nominally 10 m to 10 km in size), which may arise from the structuring of the nuclear plasma itself, or from cascading structure from TIDs. Cascading ionization structure appears naturally in the post-sunset equatorial ionosphere and in the polar-auroral ionosphere by various instability processes that I will not discuss here.

The second task I was assigned at MRC was to analyze channel probe scattering function data, collected by Dr. Roy Basler of SRI International on an HF propagation path from Narssarssuaq, on the southern tip of Greenland, to Thule, in the north of Greenland [6]. My task was to explain anomalous delay-Doppler power-spectral structure seen in the scattering functions measured by the channel probe. I was successful in doing so by deriving a solution to an equation for stochastic wave propagation that could be computed along ray paths defined by Haselgrove's Equations. I will discuss this solution, called the Phase-screen/Diffraction Method (PDM), in Section 3.

With the end of the Cold War around 1990, interest in massive-nuclear-attack scenarios waned, and I began looking for an application of my acquired HF-propagation and ray-tracing expertise to a system with a non-Cold-War function. At that time, three production versions of the Relocatable Over-the-Horizon Radar (ROTHR) had been purchased by the Navy, but were no longer required for their envisioned Cold War application. A counter narco-terrorism mission was instead defined for these radars, to watch for drug-smuggling planes transiting the Caribbean and eastern Pacific. However, it was discovered that these generally southern-looking radars suffered at times from intense Doppler-spread surface clutter. After giving a presentation on the Phase-screen/Diffraction Method at an Ionospheric Effects Symposium, I was approached by Dr. Jason Providakes of MITRE. He informed me that I had just explained the spread Doppler clutter (SDC) problem of over-the-horizon (OTH) radar. He introduced me to his sponsors, and soon I found myself immersed in the world of OTH radar (just in the nick of time, too, as other funding in HF propagation was drying up). Our group at Mission Research developed a simulator for spread Doppler clutter called CLEM (for Clutter Effects Model). This included the important elements of OTH radar signal processing, and a model for naturally-occurring small-scale ionization structure and a formulation for the Doppler spread caused by it. It also included ray tracing using a two-dimensional implementation of Haselgrove's Equations provided by Dr. Chris Coleman, then of the Australian Defence Science Technology Organisation [7-9]. Spread Doppler clutter and the Clutter Effects Model are discussed in Section 4.

The geolocation of radar targets, known as coordinate registration (CR), is an important function of OTH radar that relies on Haselgrove's Equations. OTH radars operate by HF skywave, using propagation modes that reflect from the ionosphere to illuminate targets at great ranges, in excess of 5000 km at times. Given the diurnal and transient behavior of the ionosphere, determining the position of an OTH target geographically to within several kilometers at such great ranges is a significant challenge. "Virtual geometry" is the application of Breit and Tuve's theorem, together with Martyn's Equivalent Path Theorem, and the well-known "Secant Law" conversion from oblique to equivalent vertical propagation frequency [10]. This can be used to provide the conversion from radar "slant space" coordinates of slant range (delay) and beam-steering direction to "ground" coordinates (ground range and azimuth or latitude and longitude). Virtual geometry works surprisingly well when the ionosphere is fairly horizontally stratified, for example in mid-latitudes away from the day/night terminator. To improve on this when the ionosphere has significant horizontal gradients, it is advantageous to perform ray tracing in a model ionosphere using Haselgrove's Equations. In Section 5, I will describe an efficient capability we developed to model the three-dimensional ionosphere in real time on a single PC-class computer, inverting OTH radar soundings of the ionosphere using Haselgrove's Equations. Ray tracing in this ionospheric model provides OTH coordinate registration while the next soundings are being collected. The method is called CREDO, for Coordinate Registration Enhancement by Dynamic Optimization.

The final application of Haselgrove's Equations to be discussed here is the mitigation of the effects of traveling ionospheric disturbances on OTH radar coordinate registration. Current OTH radars do not have sufficient bandwidth and temporal resolution in their backscatter soundings to allow real-time modeling of TIDs with sufficient accuracy to account for the transient deviations they impart on OTH radar coordinate registration. However, evidence of TIDs is often visible in the soundings as modulation of the backscatter power in delay and frequency, and sometimes slight delay-frequency modulation of the backscatter-sounding leading edge. We have found a promising candidate method for mitigating the apparent target wander caused by TIDs, which uses the ionospheric-induced Doppler shifts of backscattered surface clutter to infer range and azimuthal TID-induced ray deflections. The method employs a mutual-regression technique, where the regression coefficients are determined by ray tracing in

modeled TID ionospheres using Haselgrove's Equations. This approach is reviewed in Section 6.

Summary remarks are given in Section 7.

2. Ray Focusing and Homing

Haselgrove's Equations are an implementation of Hamilton's equations suitable for numerical integration on a computer. Haselgrove had in mind the determination of the paths of dominant energy flow of radiowaves in anisotropic media, specifically the ionosphere. In the Jones-Stephenson "Raytrace" code [5], these equations are written in Earth-centered spherical polar coordinates as

$$\dot{r} = \frac{\partial H}{\partial k_r}, \tag{1}$$

$$\dot{\theta} = \frac{1}{r}\frac{\partial H}{\partial k_\theta}, \tag{2}$$

$$\dot{\varphi} = \frac{1}{r\sin\theta}\frac{\partial H}{\partial k_\varphi}, \tag{3}$$

$$\dot{k}_r = -\frac{\partial H}{\partial r} + k_\theta\dot{\theta} + k_\varphi \sin\theta\dot{\varphi}, \tag{4}$$

$$\dot{k}_\theta = -\frac{1}{r}\frac{\partial H}{\partial \theta} - \frac{1}{r}k_\theta\dot{r} + k_\varphi\cos\theta\dot{\varphi}, \tag{5}$$

$$\dot{k}_\varphi = -\frac{1}{r\sin\theta}\frac{\partial H}{\partial \varphi} - \frac{1}{r}k_\varphi\dot{r} - k_\varphi\frac{\cos\theta}{\sin\theta}\dot{\theta}. \tag{6}$$

Here, (r,θ,ϕ) are the usual spherical polar coordinates, and (k_r,k_θ,k_ϕ) are the wave vector components in the $(\hat{r},\hat{\theta},\hat{\varphi})$ directions, with $k^2 = k_r^2 + k_\theta^2 + k_\varphi^2$. Additional time-frequency equations can be included for time-dependent ionospheres. The dot refers to differentiation with respect to any parameter that varies monotonically along the ray path. The Hamiltonian can be taken as any of a number of possible expressions of the dispersion relation [1], for example,

$$H = \frac{1}{2}\left[\frac{c^2}{\omega^2}k^2 - \mu^2\right], \tag{7}$$

where μ is the ionospheric index of refraction, here taken to be real (i.e., neglecting absorption). The index of refraction can be spatially varying and generally anisotropic, like the ionosphere. That is, in general $\mu = \mu\left(r,\theta,\phi,k_r,k_\theta,k_\phi\right)$.

Haselgrove's Equations trace the paths of dominant energy flow. This can be understood in the following way. Consider all possible paths connecting two points. Those paths about which small spatial path deviations cause significant phase changes cannot efficiently propagate energy: significant phase changes result in destructive interference and do not allow energy transport. Only those paths with stationary phase for which small spatial deviations result in little or no phase change will have nearby "micro-rays" arriving in phase, interfering constructively and allowing significant levels of energy transport. Hamilton's equations are an expression of the principle of stationary phase, being derived by the calculus of variations from Fermat's principle of stationary phase,

$$0 = \delta\int_a^b \vec{k}\bullet d\vec{s}\,. \tag{8}$$

Here, δ represents any spatial variation of a path between points *a* and *b*, with the endpoints remaining fixed, and the integral is over the path. Any equations derived from Equation (8) using the variational calculus, as are Equations (1)-(6), will embody phase stationarity. Hence, Haselgrove's Equations describe the paths of dominant energy flow in an ionosphere with spatially varying and generally anisotropic index of refraction. These stationary phase paths are referred to as "rays," and the numerical solution of the Haselgrove Equations is referred to as "ray tracing."

The equations for an infinitesimal flux tube about a given ray can be derived using secondary variations of Haselgrove's Equations. I did this for the isotropic ionosphere in [11]. Bennett [12] cited other independent discoveries of similar methods. The generalization to anisotropic ionospheres was developed by Västberg and Lundborg [13]. Using the example of the spherical polar coordinate expression of Haselgrove's Equations in Equations (1)-(6), the idea is simply to vary these equations with respect to the independent variables (r,θ,ϕ) and (k_r,k_θ,k_ϕ) by differentiation to develop equations for $\delta\dot{r},\delta\dot{\theta},\delta\dot{\varphi},\delta\dot{k}_r,\delta\dot{k}_\theta,\delta\dot{k}_\varphi$ in terms of infinitesimal deviations $\delta r,\delta\theta,\delta\varphi,\delta k_r,\delta k_\theta,\delta k_\varphi$. I will not repeat the equations here, as the expressions are fairly long, but refer the reader to [11]. These equations, like Haselgrove's Equations, are suitable for numerical integration on a computer. The integration can be initiated with deviations corresponding to infinitesimal elevation and azimuth offsets from a primary ray direction, with the integrated result giving the infinitesimally separated positions and directions of the offset rays from the primary ray at the end of the ray path. Key to the success of this formulation is that when power density is calculated (power per unit area), the size of

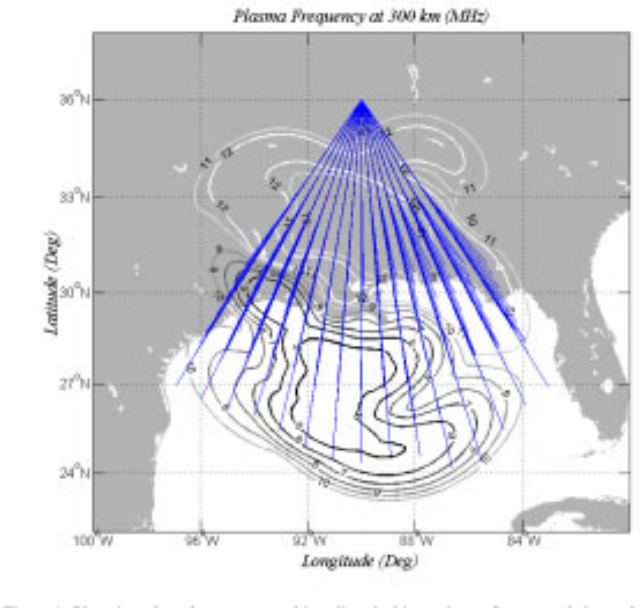

Figure 1. A plan view of ray fans propagated in a disturbed ionosphere. The contours of plasma frequency in MHz are shown.

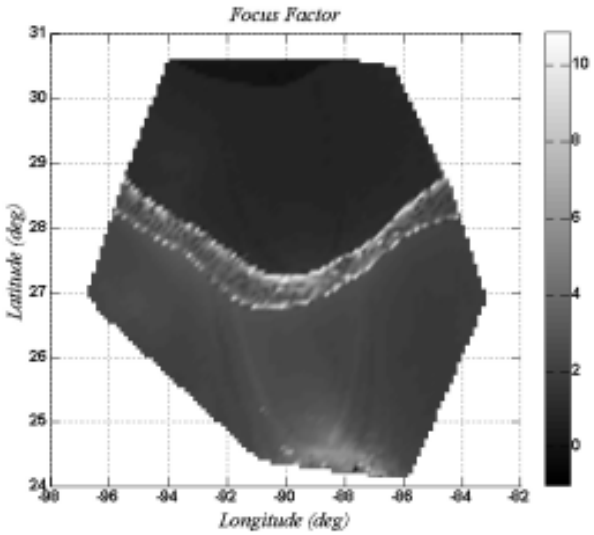

Figure 2. The focus factor for a dense set of rays (0.1° separation in azimuth and elevation), corresponding to the sparse set of rays shown in Figure 1

the initial (infinitesimal) deviations divides out of the problem, allowing a truly differential calculation to be performed: the equations describe a purely differential flux tube.

The utility of "single-ray focusing" calculations is illustrated in Figures 1 and 2. Figure 1 shows a sparse set of 10 MHz rays propagated in a disturbed ionosphere; contours of plasma frequency at 300 km altitude are shown (in MHz). The ray fans in Figure 1 were separated by 5° in azimuth, and within each fan the rays were separated by 2° in elevation, beginning at 8° and ending at 30°. Figure 2 displays the focus factor for a much denser set of rays (separated by 0.1° in azimuth and elevation). Focus factor is defined as the ratio of received power to that expected for purely $1/s^2$ divergence, where s is the geometric path length. In Figure 2, the focus factor is plotted on a linear scale, and I artificially limited the focus factor to a maximum value of 10, an arbitrary but sensible value to account for the fact that extreme focusing will not happen in actuality: wave-interference effects that are not included in stationary-phase ray theory limit the amount of focusing that can actually occur at caustics (where the cross sectional area of a flux tube collapses to zero). Two such caustics are evident in Figure 2 as the two white lines running across azimuth. Note that there were plotting artifacts in Figure 2, due to the nonuniform laying down of the finitely-separated ray landing points, and the necessity of interpolating to a uniform grid for the functioning of the contour-plotting algorithm. The caustic lines should actually be continuous in Figure 2.

The equations for infinitesimally-deviated rays used in the single-ray focusing calculations can also be used in a very efficient ray-homing algorithm. The couplet from Longfellow's famous poem, "I shot an arrow into the air, It fell to Earth, I knew not where," is quite true of ray tracing. One can specify the initial conditions of a ray, its launch direction in particular, but where that ray lands is generally not known until after the calculation. Yet, it is often desired to employ ray tracing in the study of fixed links, and determining the initial launch direction that will land the ray at the desired link endpoint position via human-in-the-loop iteration can be a time-consuming and frustrating occupation, especially if the ionosphere has nontrivial three-dimensional structure. We have developed an automated algorithm for ray homing that makes use of the information about infinitesimally deviated rays used in the ray-focusing calculation.

Consider the Earth-centered spherical polar coordinate system of Equations (1)-(6). Let α represent the ray launch azimuth measured clockwise from north, and let β represent the ray launch elevation. Assume the rays of interest land on the horizontal Earth's surface. The change in ray landing point with changes in launch azimuth and elevation are given by

$$\frac{\partial\theta}{\partial\alpha} = -\frac{\partial\theta}{\partial\varphi_k} + \frac{1}{r}\frac{k_\theta}{k_r}\frac{\partial r}{\partial\varphi_k}, \tag{9}$$

$$\frac{\partial\theta}{\partial\beta} = -\frac{\partial\theta}{\partial\theta_k} + \frac{1}{r}\frac{k_\theta}{k_r}\frac{\partial r}{\partial\theta_k}, \tag{10}$$

$$\frac{\partial\varphi}{\partial\alpha} = -\frac{\partial\varphi}{\partial\varphi_k} + \frac{1}{r\sin\theta}\frac{k_\varphi}{k_r}\frac{\partial r}{\partial\varphi_k}, \tag{11}$$

$$\frac{\partial\varphi}{\partial\beta} = -\frac{\partial\varphi}{\partial\theta_k} + \frac{1}{r\sin\theta}\frac{k_\varphi}{k_r}\frac{\partial r}{\partial\theta_k}, \qquad (12)$$

where θ_k and φ_k are the spherical polar coordinates of **k** with respect to $\hat{r}, \hat{\theta}, \hat{\varphi}$ as local Cartesian axes. The automated homing algorithm proceeds by shooting a ray, computing the miss distances to the desired homing point ($\Delta\theta, \Delta\varphi$), and adjusting the ray-launch azimuth and elevation by

$$\Delta\alpha = \frac{\frac{\partial\theta}{\partial\beta}\Delta\varphi - \frac{\partial\varphi}{\partial\beta}\Delta\theta}{\frac{\partial\theta}{\partial\beta}\frac{\partial\varphi}{\partial\alpha} - \frac{\partial\theta}{\partial\alpha}\frac{\partial\varphi}{\partial\beta}}, \qquad (13)$$

$$\Delta\beta = \frac{\Delta\theta - \frac{\partial\theta}{\partial\alpha}\Delta\alpha}{\frac{\partial\theta}{\partial\beta}}, \qquad (14)$$

and then iterating the procedure until the ray lands within some specified tolerance from the desired landing point.

An example of the performance of the automated ray-homing algorithm is shown in Figure 3. In this case, we used the ionosphere represented in Figure 1, and sought a ray connecting latitude 36°N and longitude 90°W to a homing point at latitude 30°N and longitude 85.6°W. The initial ray (10 MHz) was launched at an azimuth of 155° from north and at an elevation of 20°. The convergence tolerance was set to get the final ray within 0.01 km of the desired homing point. Eight iterations were required, taking only about a tenth of a second on my 2.0 GHz laptop. The final azimuth and elevation angles were 149.352316° and 25.637033°, respectively. In this case, as is typical, the algorithm found the low ray mode nearest the initial launch direction that reached the desired homing point.

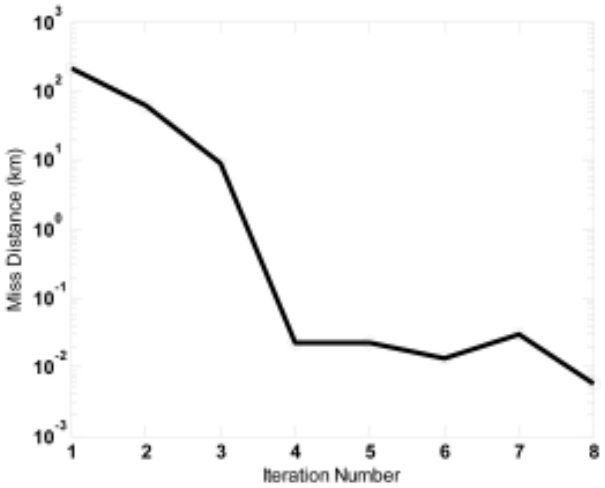

Figure 3. The miss distance as a function of the iteration number, using the automated ray-homing algorithm.

The automated ray-homing algorithm has also been incorporated in the ray-tracing package *TRACKER*, developed at Los Alamos National Laboratory [14].

3. HF Channel Modeling

During the Cold War, it was assumed that in the case of a massive nuclear attack, the military might have to rely on HF communications. However, the HF ionospheric communication channel would then be highly disturbed, and it was necessary to estimate this level of disturbance so that modems could be hardened by appropriate coding schemes that would be resilient to the channel disruptions. Small-scale ionization structure will cause a propagated HF signal to suffer angular scattering, so that received signals are spread over a range of angles and delays, owing to the extra delay of the scattered "micro-multipath" signal elements, or "micro-rays." Moving ionization (or moving link geometry through ionization) will introduce corresponding spread in signal Doppler frequency due to the spread in impinging angles of "micro-rays" on the small-scale ionization structures causing the scatter. The spreads in signal delay and Doppler are characterized by the scattering function, the delay-Doppler power spectrum of the received signal.

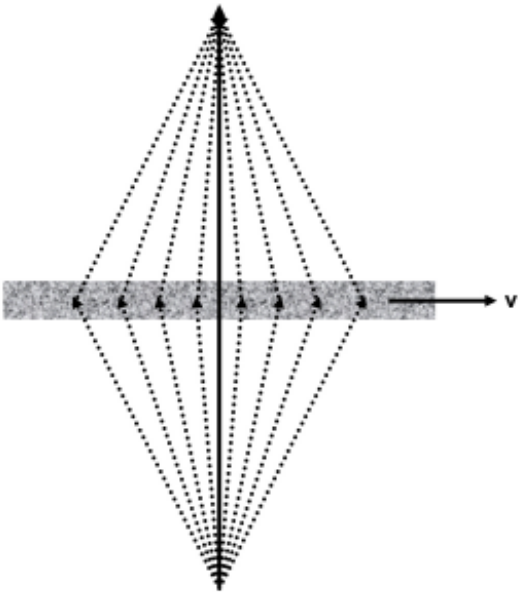

Figure 4. A plan view of a primary ray path (solid arrow) and associated micro-rays (dotted lines), scattered by an intervening phase-changing screen representing small-scale ionization structure. The phase screen was moving transversely to the primary ray with speed v.

The natural polar ionosphere develops a similar structuring of ionization at small scales to that expected in the nuclear environment anticipated by Cold War analysts. Measurements of the channel scattering function for a polar HF propagation link were funded by the US Defense Nuclear Agency and carried out by SRI International, using an HF channel probe [6]. The measurements revealed delay-Doppler correlation characteristics that were not anticipated, and I was assigned the task of explaining the measurements.

The expected delay-Doppler correlation was parabolic in shape. This expectation was derived from the following simple model. In a typical, highly oblique, HF-propagation geometry corresponding to a communication link to a remote location, much of the propagation path is in free space with significant interaction with the ionosphere confined to a region near the reflection point. Thus, it was thought that a good propagation model for scatter caused by small-scale ionization structure would be a single phase-changing screen near the midpoint of the propagation path, oriented orthogonal to that path. This is illustrated in Figure 4, which shows a plan view of the modeled path. In the absence of small-scale ionization structure, the only path connecting the transmitter and receiver would be the primary ray path. However, the presence of small-scale ionization structure allows some of the energy propagating at other angles to scatter back to the receiver in "dogleg paths." These dogleg paths necessarily have longer propagation delays than the primary ray path (ignoring refraction of the primary ray path in the vertical plane, a good assumption for highly oblique geometries), and therefore the angular spread in received micro-rays was expected to give a similar spread in received signal delay. If there is relative motion between the phase screen and the primary ray path, Doppler shifts can be imparted to the micro-rays. To first order, motion along the primary ray direction will not impart a Doppler shift; only motion transverse to the primary ray direction is significant to the Doppler-shift calculation. The wider the angle of the dogleg micro-ray relative to the primary ray direction, the larger in magnitude will be the Doppler shift imparted by the plasma structure. Those micro-rays directed against the plasma motion will be shifted up in Doppler frequency, and those directed with the plasma flow will be shifted down in Doppler frequency. This results in the aforementioned parabolic correlation of Doppler shift with delay in this model.

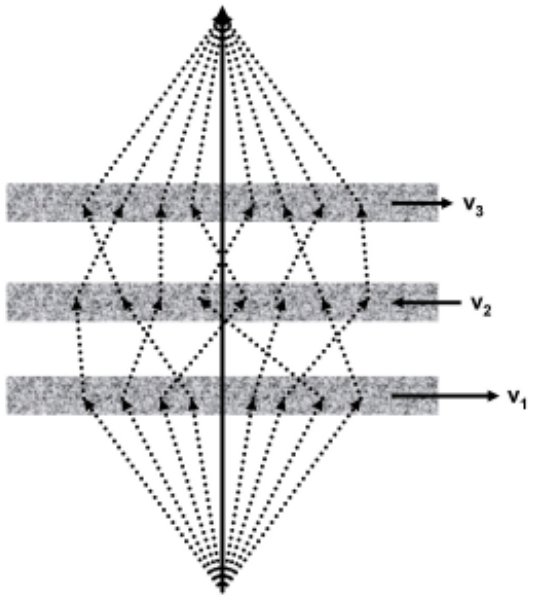

Figure 5. A plan view of a primary ray path (solid arrow) and associated micro-rays (dotted lines), scattered by three intervening phase-changing screens representing small-scale ionization structure. The phase screens were moving transversely to the primary ray with speeds v_1, v_2, and v_3.

The HF channel-probe measurements revealed scattering functions with a broad range of delay-Doppler shapes. Only rarely did the anticipated parabolic shape manifest itself. More typically, the shapes exhibited little correlation between delay and Doppler frequency. I realized this must be caused by multiple scattering over the extended range of the ray paths in the ionosphere, and that varying plasma flow structure over that distance would tend to decorrelate delay and Doppler frequency. This is illustrated in Figure 5, which again shows a plan view of the modeled propagation path, similar to Figure 4, but in this case for three intervening phase screens. Possible scatter geometries for micro-rays are drawn. Clearly, the length (delay) of the scattered micro-rays is no longer exactly correlated with their angle of arrival, nor will their Doppler shifts necessarily be correlated with their delay, especially given that the phase screens may have different speeds. The single-phase-screen model was unable to account for these extended media effects, so I set about deriving a multiple-phase-screen theory.

The phase-screen approach to stochastic wave propagation solves the parabolic wave equation (PWE) for either the scattered wave itself, or for the mutual coherence function of that wave. The mutual coherence function is the correlation function of the received signal over separated frequencies, positions, and times. The scattering function is related to the mutual coherence function by Fourier transforms in each dimension. The multiple-phase-screen approach to the numerical solution of the parabolic wave equation had been worked out by MRC's Dr. Dennis Knepp [15]. Knepp also derived an analytical solution for the two-frequency mutual coherence function, for the special cases of an extended region of structured ionization upon a uniform background, where the structure is either highly elongated or isotropic [16]. These works served as a guide to my approach.

Solving the parabolic wave equation using multiple phase screens proceeds by imparting phase changes to the incident wave at the first screen, then allowing diffraction to develop using Huygens-Fresnel propagation to the next

screen (which is implemented in a simple Fourier-transform procedure), and repeating the process through all the screens and then to the receiver. This procedure is now usually called the Fourier split-step method, but I referred to it at the time as the Phase-screen/Diffraction Method (PDM). My contribution to the theory was the realization that the Phase-screen/Diffraction Method solution approach, when applied to the parabolic wave equation for the mutual coherence function (as opposed to the parabolic wave equation for the wave itself), could be solved analytically for the general case of nonuniform background ionization and varying screen velocities, in the strong-scatter regime [17]. I implemented this solution on primary rays computed using Haselgrove's Equations in various ionosphere models, including ionosphere models derived from ionograms and incoherent-scatter radar measurements taken during the course of the SRI channel-probe campaigns. These results clearly showed that the unexpected scattering-function shapes, measured by the channel probe, were due to nonuniform plasma motion of the polar ionosphere along the extended length of the HF ray paths of the channel-probe link.

Examples of Phase-screen/Diffraction Method-derived scattering functions are shown in Figure 6. For these cases, Haselgrove's Equations were used to propagate a 12 MHz ray from the southern tip of Greenland to a point approximately 1464 km to the north on a one-hop F-region low-ray propagation mode in a generic, but strong, daytime-ionosphere model. The traced ray determined the encountered electron density along the path, as well as the angles of the ray direction with respect to the geomagnetic field. At this high-latitude location and for a northerly directed and highly oblique path like this, the geomagnetic field lines crossed the ray path almost orthogonally, vertically. Naturally occurring small-scale ionization irregularity structure tends to be elongated along the geomagnetic field lines, with scale sizes that are reasonably parameterized by a power spectral density (PSD) with an outer scale size of $L_{\perp} = 10$ km orthogonal to the geomagnetic field lines and $L_{\parallel} = 100$ km along them, and a so-called freezing scale of $l_f = 1$km, where is assumed that the power spectral density breaks to a steeper spectral slope, indicating declining presence of relevant structure. (Details about the power spectral density parameterization can be found in [17]. For the interested reader, in the examples of Figure 6, using the notation of [17], I used a two-component power spectral density, with an inner spectral slope of $n = 1.9$, breaking to the slightly steeper spectral slope of $n' = 2$ at the freezing scale. The strength of the power spectral density of the irregularity structure was taken to be 20% of the ambient electron density at each point along the ray, that is, $\sigma_{Ne} / \langle N_e \rangle = 0.2$.) To generate the Phase-screen/Diffraction Method solution for the scattering function, I collapsed the ionization to 11 or 12 equidistant phase screens along the ray path, and applied the average aspect angle of the ray to the geomagnetic field lines over each collapsed interval.

In Figure 6a, all 12 phase screens were moving together in the positive (easterly) direction, resulting in a scattering function that exhibited the parabolic delay-Doppler correlation that had been anticipated by the simple

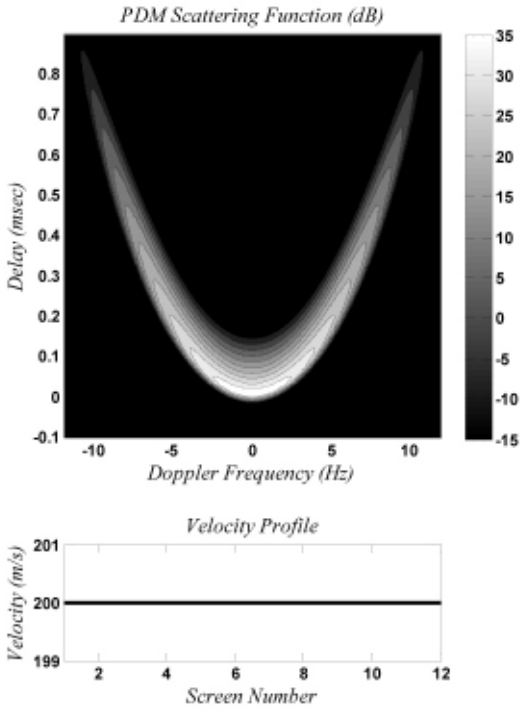

Figure 6a. Phase-screen/Diffraction Method scattering functions for uniform plasma motion.

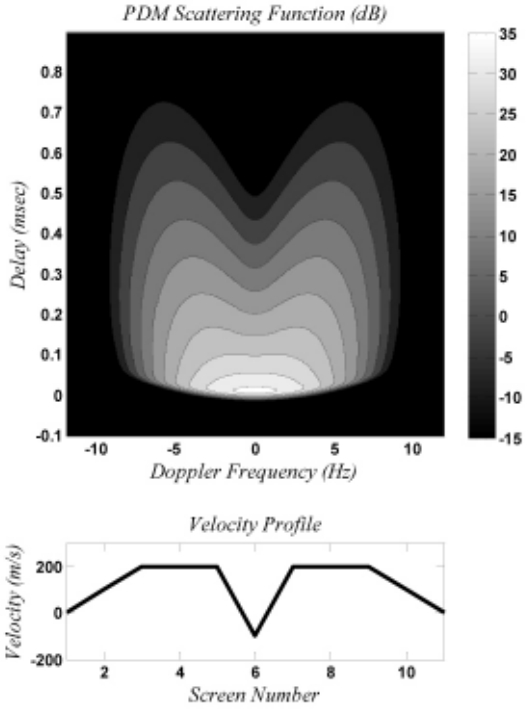

Figure 6b. Phase-screen/Diffraction Method scattering functions for a dominantly positively moving plasma with a negative velocity sheer.

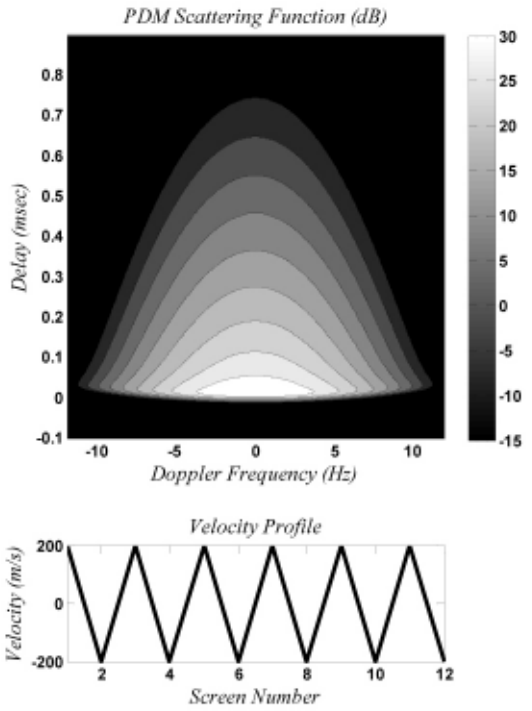

Figure 6c. Phase-screen/Diffraction Method scattering functions for turbulent plasma motion

single-phase-screen model of Figure 4, except that scatter occurring over the extended region represented by the multiple screens caused a slight decorrelation of delay and Doppler that manifested as a slower decay in delay than rise at each Doppler frequency. The single-phase-screen result would have had symmetric rise and fall times at each Doppler frequency. In Figure 6b, the plasma was assumed to be moving in a dominantly positive direction, except for a narrow region near the center of the ray path that had a strong negative (westerly) flow, representing a plasma velocity shear. The result was a scattering function that retained some of the wing-like structure of Figure 6a, but with a substantial loss of delay-Doppler correlation. Figure 6c was an approximation to a turbulent ionosphere with rapidly changing plasma motion along the ray path. Here, all correlation of delay with Doppler was lost.

The range of scattering-function shapes and their degree of spreading in delay and Doppler (and angle-of-arrival) is of interest to designers of communication modems and radars that must include algorithms to mitigate the deleterious effects of this spreading on, for example, inter-symbol interference or radar resolution. Using the approach of modeling disturbed ionospheres, applying Haselgrove's Equations to trace rays, and computing the implied scattering functions, our team at MRC was able to develop an HF signal specification for nuclear-disturbed environments to assist such mitigation-algorithm development.

4. Spread Doppler Clutter

Over-the-horizon radars generally operate as Doppler radars, that is, they transmit a series of pulses or linear FM "chirps" that can be coherently integrated and Doppler processed to separate moving targets from the inevitable strong surface return. This surface return is either ground clutter at zero Doppler shift, if the return is from land, or Bragg clutter, if the return is from the ocean's surface. Bragg clutter results from the coherent addition of signal elements backscattering from ocean-wave components that are separated by half the radar's operating wavelength, and is usually manifested as two clutter peaks shifted slightly above and below zero Doppler frequency by an amount corresponding to the speed of the ocean's waves toward and away from the radar. The Bragg peaks will be more or less resolved depending on the OTH radar's coherent integration time. In either case (land or sea returns), moving targets with a significant velocity component toward or away from the radar can be separated from the surface clutter by Doppler processing, as long as the surface clutter remains narrowly distributed in Doppler frequency.

As I mentioned in the introduction, I was approached by Dr. Jason Providakes of MITRE following a talk on the Phase-screen/Diffraction Method that I had given at the 1990 Ionospheric Effects Symposium. He apprised me of what has come to be called the spread Doppler clutter problem of OTH radar. Later, at the Naval Research Laboratory, Joe Thomason and Ben Root showed me some examples of spread Doppler clutter (SDC) measured by the prototype ROTHR system in Virginia, which looked in southerly directions. Post-sunset on many nights, the surface clutter was observed to broaden in Doppler frequency, sometimes to the point that it completely filled the ambiguous Doppler window (Doppler processing of OTH radar returns causes Doppler frequencies with magnitudes larger than the Nyquist frequency to wrap back into the first Nyquist interval). Radar targets at any Doppler shift could be obscured. It was a significant and unanticipated problem.

I knew that often post-sunset low-latitude ionospheric structure develops in equatorial plasma bubbles. In 1988, I had been involved in a campaign to measure scintillation using the ALTAIR radar, viewing satellites through this structure. I asked Joe and Ben about OTH signal processing, and was told about the range ambiguity of chirped waveforms, and that returns from ranges beyond the first ambiguity region appear superimposed on the returns in the first region. It seemed reasonable to me that the spread Doppler clutter was being caused by multi-hop rays that were making their way to the low-latitude ionosphere, and were being phase-modulated by the structured ionization there.

The Australians were seeing the same spread Doppler clutter with their experimental OTH radar, called the Jindalee Facility – Alice Springs (JFAS), which has a northwesterly-

directed boresight and, similar to the ROTHR in Virginia, can look toward the equatorial zone. At the dinner for an OTH radar meeting at Hanscom Air Force Base, around hors d'oeuvres, I became involved in a conversation about spread Doppler clutter with Dr. Jason Providakes (MITRE) and Dr. Bruce Ward of the Australian Defence Science Technology Organisation (DSTO). The first ROTHR production system had just been fielded on the island of Amchitka in the Aleutians, with a boresight that allowed illumination of the Australian JFAS radar. Likewise, the JFAS radar, if steered to its easternmost extent, could just barely illuminate the Amchitka ROTHR. Bruce and Jason hit upon the idea for a one-way/two-way experiment, where the signals from each opposing radar would be received by the other radar as well as itself. To make a long story short, the experiment was conducted, and the result conclusively proved that spread Doppler clutter was being caused by phase-modulated ground clutter from propagation modes transiting the structured low-latitude ionosphere.

We at MRC set about developing a capability to simulate spread Doppler clutter, and to provide a predictive capability that could be used to assist OTH radar site selection or to predict its effects over a solar cycle. A parallel spread-Doppler-clutter simulation capability was pursued by Dr. Chris Coleman at DSTO [7]. We called our spread Doppler clutter simulator CLEM, for Clutter Effects Model [18]. Clutter Effects Model uses WBMOD, a climatological model for the occurrence of small-scale ionization structure [19, 20] to define the geographical coverage and strength of the scattering regions as a function of time of day, season, and solar cycle. Haselgrove ray tracing is used to compute the OTH radar propagation paths in a climatological ionosphere model. Since the simulation requires extensive ray-tracing calculations, we made use of an extremely fast – albeit two-dimensional – ray-tracing program provided by Dr. Chris Coleman [8].

The Clutter Effects Model includes all the radar signal processing required to simulate OTH radar clutter returns, including range and Doppler folding, ray focusing and absorption, and a ground/sea map to allow different forward/backscatter losses to be applied. We also included geomagnetic field-aligned backscatter, because small-scale ionization structure tends to stretch out along the geomagnetic field, specular scatter from the elongated structure is coherent, and if the illumination becomes orthogonal to that structure, then very bright backscatter clutter occurs. This field-aligned backscatter clutter was well known to the operators of the OTH-B system when looking northward through the polar ionosphere, since orthogonal illumination of the near-vertical field-aligned polar ionization structure was easily achieved for their highly oblique ray paths. Specular coherent scatter was also observed in the one-way/two-way JFAS-ROTHR experiment from field-aligned equatorial ionization structure. As the beams of the two radars were swept across longitude, a bright clutter return was observed on the one-way paths when the scattering angle between the two radars became specular. A surprising thing, first noted by Dr. Chris Coleman, was that some rays from mid-latitude equatorial-looking OTH radars could achieve orthogonality with the low-latitude geomagnetic field. The dip angle of the geomagnetic field in the region of the Appleton Anomalies was such that certain rays on certain hops, on their way back down from an ionospheric reflection, could impinge orthogonally on geomagnetic field-aligned structure that might be present. At Coleman's encouragement, we included the calculation of such coherent backscatter clutter in the Clutter Effects Model by monitoring the aspect angle of our Haselgrove rays with the geomagnetic field and flagging near orthogonality when it occurred, then computing the Born backscatter cross section on WBMOD-generated structure that might be there.

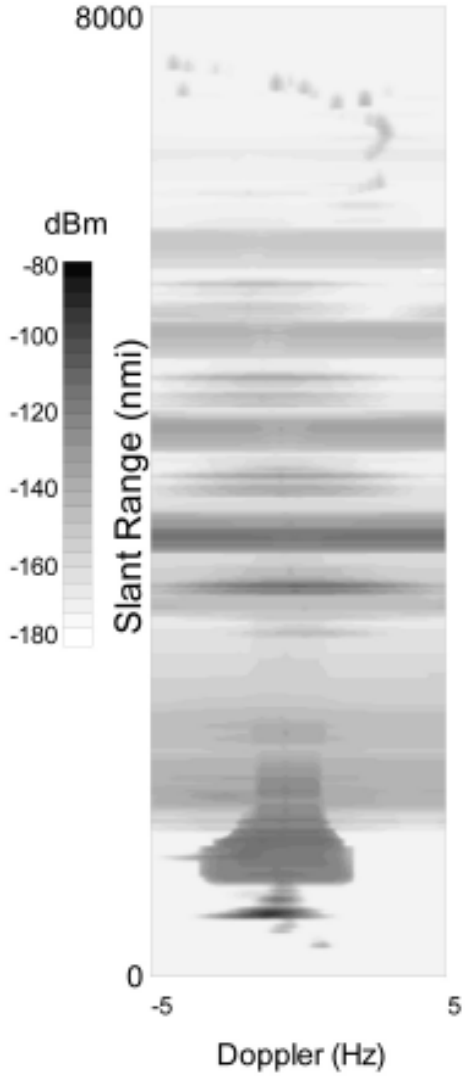

Figure 7. An amplitude-range-Doppler plot of CLEM-predicted spread Doppler clutter for a hypothetical southerly-directed, northern-latitude OTH radar (after Lauer et al. [18, Plate 4]).

An example of Clutter Effects Model output for strong spread Doppler clutter conditions is shown in Figure 7. This is an amplitude-range-Doppler (ARD) plot of predicted surface clutter returns for a hypothetical southerly-looking northern-latitude OTH radar, with an unambiguous range of 8000 nautical miles and ±5 Hz in Doppler frequency. Using typical OTH radar operating parameters, the distant clutter would be range-folded atop the nearer ranges, obstructing visibility of targets in the operational coverage area.

5. Coordinate Registration for Over-the-Horizon Radar

Skywave OTH radars, which operate in the HF band, are probably the most efficient wide-area surveillance sensor in terms of coverage area per cost. OTH radars with linear transmitting and receiving arrays (such as the US Navy's ROTHR systems) can often see out to ranges exceeding 5000 km with azimuthal coverage of 80°. These long surveillance ranges are possible because of HF skywave propagation, but the dependence of skywave-propagation-mode geometry on the particulars of the constantly-changing ionosphere make determining the geographical positions of observed targets challenging. This geo-positioning is referred to as coordinate registration (CR).

Attaining accurate coordinate registration requires accurate and timely ionosphere models, through which Haselgrove's Equations can be used to trace rays to determine the relationship between the radar-measured slant-range (group-path length, i.e., delay times the speed of light) and beam-steering directions and the geographic position of targets in terms of, for example, ground range and azimuth from the OTH receiver. My HF propagation group at MRC (now at NorthWest Research Associates) developed a capability for inverting OTH radar soundings in real time to obtain an accurate model of the current three-dimensional electron-density distribution of the ionosphere, all on a single PC-class computer. This capability is called CREDO, for Coordinate Registration Enhancement by Dynamic Optimization. CREDO makes substantial use of ray tracing using Haselgrove's Equations.

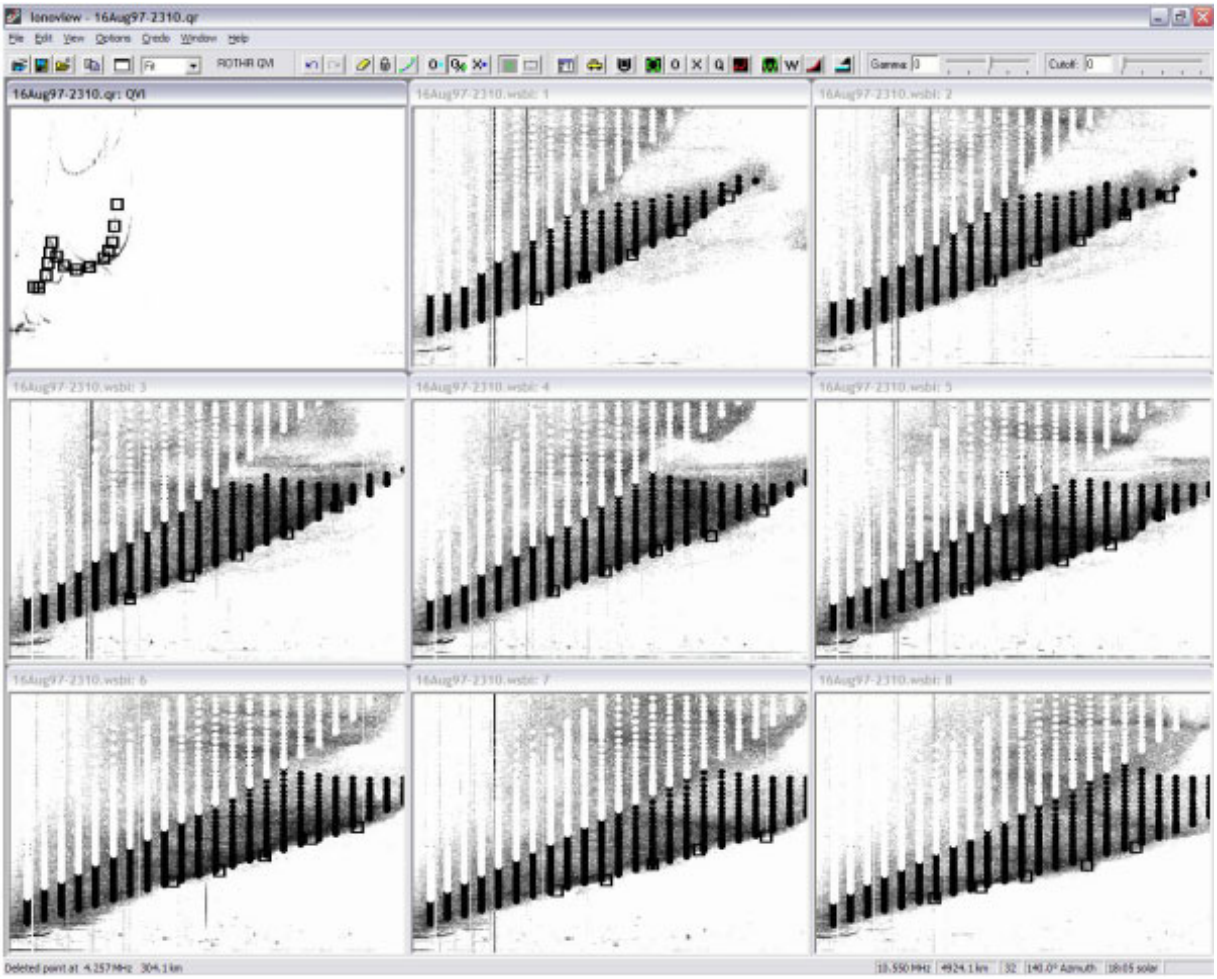

Figure 8. A CREDO display of the vertical-incidence ionogram image (top left panel), and backscatter-ionogram images, for all eight azimuthal beams of an OTH radar sounder measured on August 16, 1997, at 2310 UT. Selected vertical-incidence-ionogram ordinary trace and backscatter-ionogram leading-edge points are shown by boxes. The domain of each backscatter ionogram extended from 5 to 28 MHz (horizontal axis) and from 0 to 4940 km (one-way group path, vertical axis). The dots represent the synthesized backscatter ionograms for the reconstructed ionosphere (black for one-hop and white for two-hop propagation) (after Fridman and Nickisch [22, Plate 1]).

Typical OTH radar ionospheric soundings consist of one or more vertical-incidence ionograms (VIs) and backscatter ionograms (BIs), with soundings updated every few minutes. The backscatter ionograms are characterized by the leading edge of one-hop propagation at the edge of the skip zone as a function of sounding frequency. In the CREDO approach, we first obtain the vertical electron density over the vertical ionosonde location(s) by one of two methods. The CREDO operator can select the ordinary (o-ray) or extraordinary (x-ray) one-hop trace using a few mouse clicks, and then CREDO proceeds with direct inversion of the vertical-incidence ionogram using the standard method (augmented by a smart E-F valley model). Alternatively, in CREDO's automated mode, an algorithm automatically selects o-ray and x-ray points along the one-hop traces, and a vertical electron-density profile is adjusted until the corresponding synthesized ionogram suitably passes through the automatically-selected points. The vertical electron-density profile over the (one or more) ionosonde locations is then used to sensibly adjust a climatological three-dimensional ionosphere model so that it matches in the ionosonde locations and extrapolates reasonably to remote locations. CREDO then proceeds by selecting several backscatter ionogram leading edge points in all backscatter ionogram azimuthal sounding sectors. This selection can either be done automatically using an impressively robust automated point extraction (APEX) algorithm, or directly by operator mouse clicks. The selected backscatter ionogram data points are then fed to a three-dimensional ionospheric inversion algorithm that repeatedly adjusts the three-dimensional ionosphere model until the synthesized backscatter ionogram leading edges obtained by Haselgrove ray tracing agree with the input data. This inversion algorithm utilizes Tikhonov regularization to ensure that the resultant three-dimensional ionosphere model is the smoothest one consistent with the input data. Note that all backscatter ionogram azimuthal sectors are included in a single inversion to produce a true three-dimensional ionosphere model. All of this can be done within about a minute after the soundings are updated using a modern PC-class computer, providing a timely and accurate ionosphere model for OTH radar coordinate registration. Details on the CREDO algorithm are given in [21, 22]. CREDO also includes the ability to optionally include two-hop backscatter ionogram leading edges and backscatter-ionogram-lobe trailing-edge data.

Figure 8 shows the CREDO graphical user interface, called *Ionoview*, for an example of an OTH radar with a single vertical-incidence ionogram sounding and eight azimuthal backscatter ionogram sounding sectors. The vertical-incidence ionogram sounding is shown in the upper-left corner pane, with the remaining panes exhibiting the backscatter ionogram soundings, each separated by eight degrees azimuth from the next. Boxes indicate the selected vertical-incidence ionogram trace points and backscatter ionogram leading edge data to be used in the CREDO inversion. Overlaying the backscatter ionogram soundings are the synthesized one-hop (black circles) and two-hop (white circles) slant ranges (group path lengths) of the Haselgrove-generated rays in the resultant CREDO ionosphere model fit, providing confirmation of the accuracy of the inversion.

In our development of CREDO, we benefited by collaboration (and friendly competition) with our colleagues at the Australian DSTO. Dr. Chris Coleman developed his own codes for solving Haselgrove's Equations, and we compared the speed of our respective codes. (Chris won. His secret: Instead of integrating Haselgrove's Equations in spherical polar coordinates, as in the Jones-Stephenson *Raytrace* program, he integrated in a Cartesian coordinate system, which greatly simplified the equations and reduced the number of computational operations required to trace a ray. He then transformed the result at the end to spherical polar coordinates.) Coleman used his Haselgrove integrator to develop a nice capability for simulating backscatter ionograms [9].

6. TID Mitigation

Traveling ionospheric disturbances (TIDs) remain a troubling cause of coordinate registration errors for OTH radar. OTH radar coordinate registration algorithms are not currently designed to account for TID-induced apparent motion, and hence large coordinate-registration errors can result. So far, it has been impossible to mitigate the effect of TID-induced motions on OTH radar return signals. Conventional OTH radar soundings have neither the spatial nor temporal resolution to resolve TIDs. Ionospherically-induced Doppler shift (IDOP) can be observed on OTH radar backscatter returns as a shifting of the centroid of the surface clutter away from zero Doppler frequency. We know that to a large extent, ionospherically-induced Doppler shift is caused by TID activity. We also know a lot about the physics of TIDs. They are driven by acoustic gravity waves (AGWs), which are buoyancy waves in the neutral atmosphere akin to water waves in the ocean. Because of the exponentially decreasing density of the neutral atmosphere with altitude, the energy from fairly minor disturbances in the lower atmosphere turns into large-amplitude waves at altitude, and these drive the ionosphere up and down the geomagnetic field lines.

In [23], we used Haselgrove ray tracing in modeled TID-affected ionospheres to demonstrate that for HF skywave paths, the slant-range rate determined by Doppler shift does not generally agree with the slant-range rate determined by the rate of change of group path, and this disagreement can be large, by factors of two or three or even more. We also showed that this disagreement appears in actual OTH radar measurements. The explanation for this discrepancy has to do with certain "path change" terms for refracted ray paths in the Fermat's Principle calculation, which cancel each other in the phase path rate calculation (Doppler shift), but add together in the calculation of group path rate. An example is shown in Figure 9. In this case, 10 MHz Haselgrove rays were homed to a fixed receiver location over a 2000 km link on an F2-layer low-ray propagation mode while a simulated TID passed through

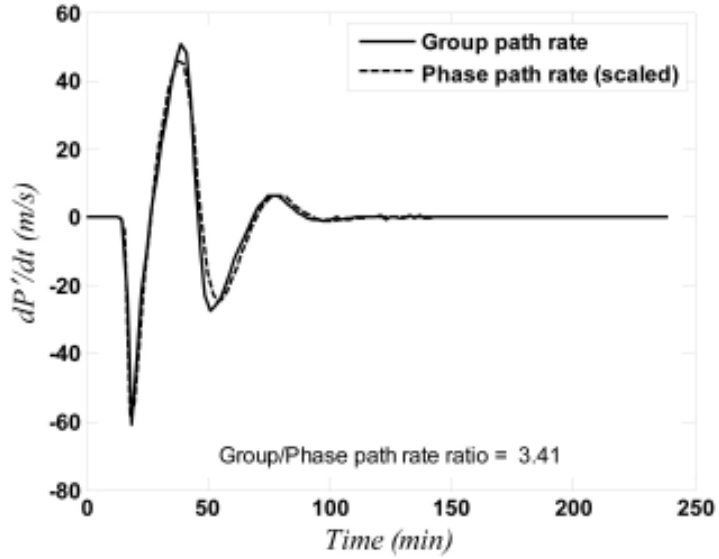

Figure 9. The scaled Doppler (dashed curve) and group path rate, (solid curve) as functions of time for an HF path to a static target affected by a simulated TID. The signal frequency was 10 MHz (from Nickisch et al. [23]).

the reflection region. Group path rate (solid curve) and phase path rate (dashed curve) were computed, and the phase path rate was scaled to provide the best least-squares fit to the group path rate curve. In this case, the phase path rate had to be multiplied by a factor of 3.41 to overlay the group path rate curve. That is to say, the apparent motion of the (static) target, induced by the TID, exhibited a discrepancy of more than a factor of three between the rates computed by group path and phase path (Doppler).

The size of this group/phase path rate discrepancy depends on various factors that were discussed in Nickisch et al. [23]. Suffice it to say that it is variable and does not always appear as a strict proportionality. However, we have found that it can be well modeled by mutual regression. That is, a time history and/or spatial sampling of ionospherically-induced Doppler shift can be used to predict group path rate, and the same technique can be used to predict apparent azimuth rate of the ray path induced by passing TIDs. We demonstrated this in [24]. Using Haselgrove ray tracing, we synthesized several HF links with passing TIDs, and generated regression coefficients for using ionospherically-induced Doppler shift to predict slant range rate (group path rate) and azimuth rate. We then used these coefficients to correct the slant range and azimuth deflections induced by other independent TID realizations. An example is shown in Figure 10, which displays TID-induced slant range deviations of a static target at a distance of 2000 km as would be measured by an OTH radar (solid curve), and the corrected result (dashed curve) obtained using the regressive prediction based on ionospherically-induced Doppler shift sampled throughout the radar's dwell illumination region. The errors have been reduced by more than 93%. In this case, no temporal regression was applied ($n_r = 0$), and the regression was a purely spatial regression.

7. Summary

Jenifer Haselgrove's Equations have been crucial to the HF propagation analyses of my group, first at Mission Research Corporation and now at NorthWest Research Associates. I have described their use in ray-focusing and

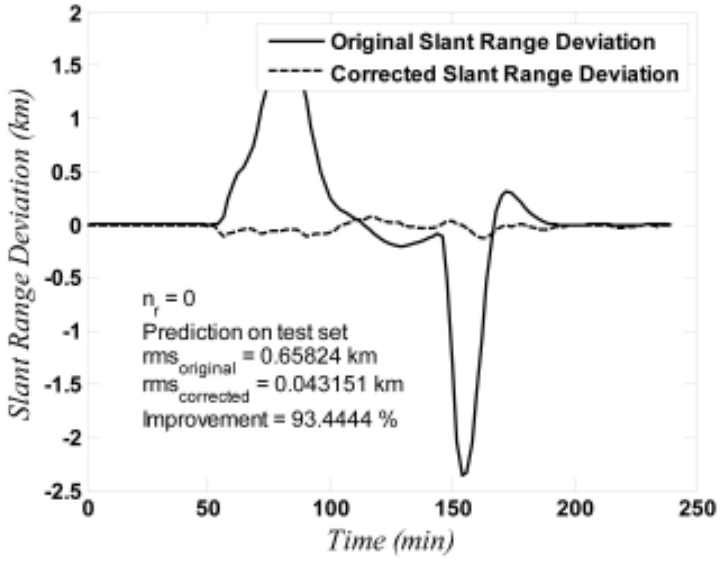

Figure 10. The slant range (group path) deviation induced on a simulated OTH radar observation of a fixed target caused by a passing TID (solid curve), and the corrected result after applying mutual-regression correction based on ionospherically-induced Doppler shift sampled throughout the OTH radar's dwell illumination region (from Nickisch et al. [24]).

ray-homing calculations for analyzing HF communication links, in HF channel modeling for disturbed ionospheres, in understanding and predicting spread Doppler clutter for OTH radar, improving coordinate registration for OTH radar, and mitigating the deleterious effects of traveling ionospheric disturbances for OTH radar. We have performed many other studies using ionospheric ray tracing, which appeared in various reports that our group produced, presentations we gave to our sponsors, or were simply tucked away in the folders of my filing cabinets and, now, the folders in my computer. I'm sure when she coded her equations on the Cambridge EDSAC back in 1954 and could trace a single ray in a mere ten minutes time, Jenifer Haselgrove had no concept that fifty years later, scientists would be running around with compact laptop computers with her equations coded in them, tracing ionospheric ray paths in milliseconds.

8. Acknowledgements

A number of colleagues at Mission Research Corporation (MRC) and NorthWest Research Associates (NWRA) were involved in the work presented here. Their names appear in the cited publications. I would particularly like to express my gratitude to Dr. Dennis Knepp, my mentor and boss for many years at MRC, who now joins me at NWRA, and to Drs. Sergey V. Fridman and Mark A. Hausman, who also joined me in the transition to NWRA following MRC's disappearance as an independent corporate entity. Our close-knit group has worked together creatively, efficiently, and compatibly for many years now. We are all friends, who understand how to work together to maximize the effectiveness of our individual talents. I would also like to acknowledge ex-MRC scientists Dr. William Wortman, and especially Mr. Carl Lauer, who had a particularly prominent role in assembling the Clutter Effects Model capability. I want to thank SRI International's Drs. Roy P. Basler and James R. Barnum, who helped me selflessly as I floundered to get established in the world of HF propagation in my early MRC years: Roy, by supplying his HF channel probe data together with hours of explaining his experimental procedures and supplying software for reading his data, and Jim, by patiently educating me about OTH radar and providing me with software and data from his WARF OTH radar to get our CREDO project off to a quick start. Finally, I would like to thank Jenifer Haselgrove for her Haselgrove Equations, which have made all this work possible, and to R. Michael Jones, for his implementation of Haselgrove's Equations into a high-fidelity and endlessly extendable ray-tracing program. I recall telling him when I once met him that I had built my career on his code.

9. References

1. K. G. Budden, *Radio Waves in the Ionosphere*, Cambridge, Cambridge University Press, 1966.
2. J. Haselgrove, "Ray Theory and a New Method for Ray Tracing," in *The Physics of the Ionosphere: Report of the Physical Society Conference on the Physics of the Ionosphere, held at Cavendish Laboratory, Cambridge, September 1954*, London, Physical Society, 1955, pp. 355-364.
3. J. Haselgrove, "Oblique Ray Paths in the Ionosphere," *Proc. Phys. Soc.*, **70B**, 1957, p. 653.
4. C. B. Haselgrove and Jenifer Haselgrove, "Twisted Ray Paths in the Ionosphere," *Proc. Phys. Soc.*, **75**, 3, March 1, 1960, pp. 357-363.
5. R. M. Jones and J. J. Stephenson, "A Versatile Three-Dimensional Ray Tracing Computer Program for Radio Waves in the Ionosphere," OT Rep. 75-76, US Dept. of Commerce, Office of Telecommunications, October 1975.
6. R. P. Basler, G. H. Price, R. T. Tsunoda, and T. L. Wong, "Ionospheric Distortion of HF Signals," *Radio Science*, **23**, 4, July-August 1988, pp. 569-579.
7. C. J. Coleman, "A Model of Doppler Spread Clutter," in J. M. Goodman (ed.), *Proceedings of the 1993 Ionospheric Effects Symposium*, Arlington VA, SRI International, 1993, pp. 37-41.
8. C. J. Coleman, "A General Purpose Ionospheric Ray Tracing Procedure," Report No. SRL0131TR, Defence Science and Technology Organisation, Australia, 1993.
9. C. J. Coleman, "On the Simulation of Backscatter Ionograms," *J. Atmos. Solar Terr. Phys.*, **59**, 1997, pp. 2089-2099.
10. Kenneth Davies, *Ionospheric Radio*, London, Peter Peregrinus Ltd., 1990.
11. L. J. Nickisch, "Focusing in the Stationary Phase Approximation," *Radio Science*, **23**, 2, March-April 1988.
12. J. A. Bennett, P. L. Dyson, R. J. Norman, "Progress in Radio Ray Tracing in the Ionosphere," *The Radio Science Bulletin*, No. 310, September 2004, p. 81.
13. A. Västberg and B. Lundborg, "Signal Intensity in the Geometrical Optics Approximation for the Magnetized Ionosphere," *Radio Science*, **31**, 6, 1996, pp. 1579-1588.
14. P. E. Argo, P. E., D. Delapp, C. D. Sutherland, and R. G. Farrer, "TRACKER: A Three-Dimensional Raytracing Program for Ionospheric Radio Propagation," NASA STI/Recon Technical Report N, V95, 24015, 1994.
15. D. L. Knepp, "Multiple Phase-Screen Calculation of the Temporal Behavior of Stochastic Waves," *Proceedings of the IEEE*, **71**, 6, June 1983, pp. 722-737.
16. D. L. Knepp, "Analytic Solution for the Two-Frequency Mutual Coherence Function for Spherical Wave Propagation," *Radio Science*, **18**, 4, July-August 1983, pp. 535-549.
17. L. J. Nickisch, "Non-Uniform Motion and Extended Media Effects on the Mutual Coherence Function: An Analytic Solution for Spaced Frequency, Position, and Time," *Radio Science*, **27**, 1, January-February 1992.
18. Carl Lauer, L. J. Nickisch, and William Wortman, "Prediction of Over-the-Horizon Radar Clutter Using the Clutter Effects Model," *Radio Science*, **33**, 4, July-August 1998.
19. J. A. Secan, "WBMOD Ionospheric Scintillation Model," NWRA-CR-93-R098, NorthWest Research Associates, Bellevue, WA, 1993.
20. J. A. Secan, R. M. Bussey, E. J. Fremouw, and S. Basu, "High Latitude Upgrade to the Wideband Ionospheric Scintillation Model, *Radio Science*, **32**, 4, 1997, pp. 1567-1574.
21. S. V. Fridman, "Reconstruction of Three-Dimensional Ionosphere from Backscatter and Vertical Ionograms Measured by Over-the-Horizon Radar," *Radio Science*, **33**, 4, 1998, pp. 1159-1171.
22. S. V. Fridman and L. J. Nickisch, "Generalization of Ionospheric Tomography on Diverse Data Sources: Reconstruction of the Three-Dimensional Ionosphere from Simultaneous Vertical Ionograms, Backscatter Ionograms, and Total Electron Content Data," *Radio Science*, **36**, 5, September-October 2001.
23. L. J. Nickisch, Mark A. Hausman, and Sergey V. Fridman, "Range Rate – Doppler Correlation for HF Propagation in Traveling Ionospheric Disturbance Environments," *Radio Science*, **41**, 2006, RS5S39, doi:10.1029/2005RS003358.
24. L. J. Nickisch, Mark A. Hausman, and Sergey Fridman, "Traveling Ionospheric Disturbance Mitigation for OTH Radar," *Proceedings IEEE Radar 2007 Conference*, 2007.

Some Phase Path Modeling Applications of Haselgrove Ray Tracing

R.I. Barnes

Abstract

Modeling ionospheric radiowave propagation, using ray-tracing techniques based on Jenifer Haselgrove's pioneering work in the 1950s, has been of significant importance in advancing HF systems such as over-the-horizon radar (OTHR). However, much of the propagation modeling for HF systems has concentrated on group path, for instance, in synthesizing the group-to-ground-range transformation space required for OTHR coordinate registration. This paper presents an investigation of a couple of ideas that exploit the ability to estimate phase path along with the group path on a ray trajectory. It is shown that a model of the phase ionosonde technique can be demonstrated using a high-precision application of Haselgrove ray tracing. This can be further extended to progress a method to resolve micro-multipath OTHR detection of targets flying above the Earth's surface.

1. Introduction

In the early to mid-1950s, a young Cambridge researcher named Jenifer Haselgrove developed a formulation describing the application of Hamiltonian principles to radiowave propagation in the ionosphere [1, 2, 3] (and later with her husband [4]). Many people were working on radio ray propagation in the 1940s and 1950s [5], but it was Haselgrove who made the significant connection between (wo)man and the then emerging machines called computers. The real innovation of Haselgrove's efforts was that it brought ionospheric propagation theory and electronic processing together, in what we now take for granted as numerical ray tracing. This has proven to be a revelation, and since then the progress in technologies that exploit ionospheric propagation have been boosted by the use of propagation modeling based upon or inspired by her pioneering work.

At about the same time as Haselgrove's early work, the US Naval Research Laboratory was developing a very rudimentary over-the-horizon radar (OTHR) system, called Multiple Storage, Integration, and Correlation (MUSIC). Both inventions have come a long way in the ensuing 50 years. In OTHR technology, ray tracing based on Haselgrove's formulation has been used in OTHR coordinate registration [6, 7, 8, 9], and in oblique- and backscatter-ionogram synthesis [10]. Barnes also showed that ray tracing could quite adequately handle estimates of polarization and Faraday rotation, even after propagation through rather severe electron-density gradients [11].

Most of these applications require relatively accurate characterization of the group path, but the phase path is of no real concern. For example, in OTHR coordinate registration, all that is required is a relatively accurate account of the group- (more accurately, radar-coordinates) to-ground location correction. Here, I wish to show some other possible applications of Haselgrove ray tracing in the HF domain that require accurate calculation of the phase path.

2. Application One: Modeling Oblique Phase Ionograms

A two-dimensional formulation of Haselgrove's equations can be found in [12], viz.,

$$\frac{dx}{d\tau} = \frac{c}{n}\left[\cos(\chi) + \sin(\chi)\frac{1}{n}\frac{\partial n}{\partial \chi}\right], \qquad (1)$$

$$\frac{dz}{d\tau} = \frac{c}{n}\left[\sin(\chi) - \cos(\chi)\frac{1}{n}\frac{\partial n}{\partial \chi}\right], \qquad (2)$$

R. I. Barnes is with the Riverside Research Institute, 1400 Key Blvd, 22209, Rosslyn, VA, USA;
Tel: + 1 (703) 908-2111; e-mail: rbarnes@.rri-usa.org.

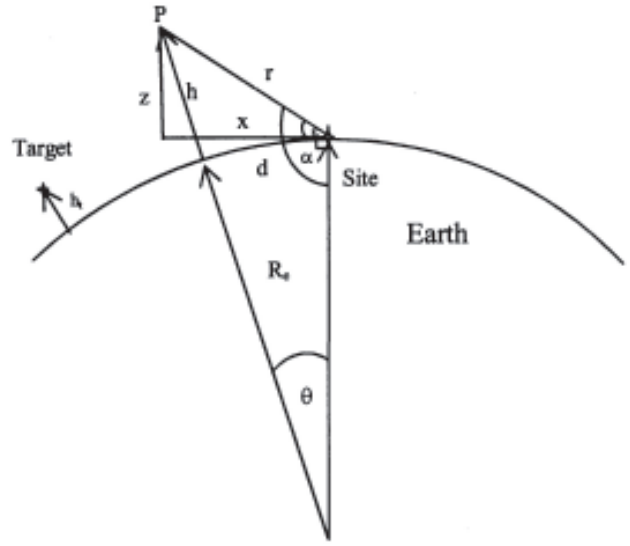

Figure 1. A Cartesian geometry consistent with the Haselgrove formulation for ray tracing set out in Equations (1)-(3). In generalizing the ray tracing for the two applications here, the receiver location has a variable height, h. For ionogram synthesis, h would normally be set to zero. The following relationships apply:

$$r = \left(x^2 + z^2\right)^{1/2},$$

$$\alpha = \pi/2 + \tan^{-1}\left(z/x\right)\left(R_e + h\right)^2 = R_e^2 + r^2 - 2rR_e\cos\left(\alpha\right)$$

$$h = \left[R_e^2 + r^2 - 2rR_e\cos\left(\alpha\right)\right]^{1/2} - R_e,$$

$$= \left[R_e^2 + r^2 - 2rR_e\left\{1/\left[1 + \left(z/x\right)^2\right]\right\}\right]^{1/2} - R_e$$

$$\frac{d\chi}{d\tau} = \frac{c}{n^2}\left[\cos\left(\chi\right)\frac{\partial n}{\partial z} + \sin\left(\chi\right)\frac{\partial n}{\partial x}\right], \qquad (3)$$

where x, z are Cartesian spatial dimensions; c is the speed of light in a vacuum; n is the refractive index at x, z; χ is the angle between the wave normal and the x axis; and τ is a parameter used to progress along the ray path.

This formulation is in Cartesian coordinates (see Figure 1), and is similar to that in [2]. These equations require solutions for the refractive index and its derivatives, and these can be supplied through the Appleton-Hartree-Lassen dispersion equation. Convenient forms are described in [13, 14], and can be readily applied to this formulation.

Haselgrove's equations provide a description of radio propagation through the ionosphere. In applying this, they may be used to synthesize ionograms through the accumulation of group range, i.e., the group delay multiplied by the speed of light in a vacuum, over ray paths. Ray tracing is an integrating process, from initial conditions, as opposed to a boundary-value solution, and so properties are accumulated as the solution progresses. In this vein, expressions for ray height, group, phase, and ground range can be added to the ray state, and integrated and stored as appropriate as the solution is progressed [15, 16]. The International Reference Ionosphere (IRI) [17] was used to provide realistic variation of the geomagnetic field and electron density, and hence refractive index, along the paths, as required by Haselgrove's equations.

The solution to the equations employed here uses a Runge-Kutta-Fehlberg (RKF) adaptive solver for efficiency (e.g., [18]), which was effectively used by Coleman [10] and Barnes [11]. The two-dimensional case was used for simplicity, and is not expected to lose generality for the current applications.

An oblique ionogram is a display of the signal strength received between fixed transmitting and receiving locations as a function of frequency and group range or delay. To accurately synthesize an oblique ionogram, the ray tracing

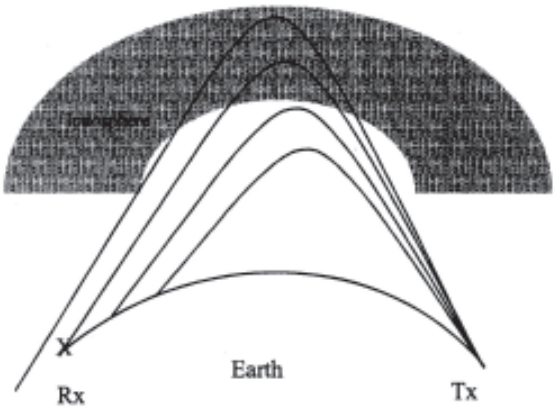

Figure 2. An oblique ionogram homing illustration. The shaded portion represents the ionosphere. Rays can be sent out at various elevations until the point of interest, X, is straddled. An iterative scheme to adjust the elevation angle based on miss distance can then be applied, resulting in a miss distance of less than 1 mm within five rays

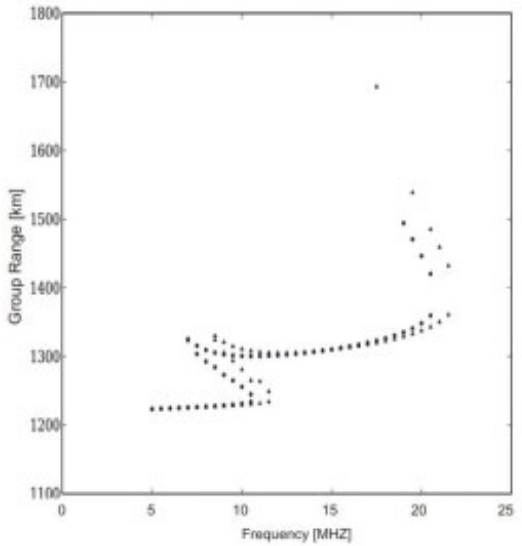

Figure 3. A simple oblique ionogram, generated using Haselgrove's equations with point-to-point homing and no power estimate.

must home in on elevation angle, to connect rays from the transmitting to receiving sites (e.g. [10, 19], and see Figure 2). To gain adequate accuracy in homing, a technique was developed using a simple linear correction to the initial ray elevation, based on the miss distances of previous rays. Once a ray is deemed close enough to the target location according to a user-defined threshold, the previous two ray states are recorded at the target height, and then interpolated to the precise target location. It was found that an error contribution to the phase range of the order of 1 mm could be achieved within five rays using this method, obviating the need for a more complex algorithm.

Sometimes, in ionogram synthesis, the relative power of paths is not needed, and the paths found to be present are simply displayed as a point or part of a line segment on the range-frequency diagram. This can be done readily with the current implementation, e.g., see Figure 3.

We now wish to advance this ionogram-generation process to recreate so-called phase ionograms [20, 21, 22]. In practice, a phase ionosonde coherently detects the incoming signal, and then uses a mechanism to represent the phase as detected on a video display unit. This has been done by using the peaks from a phase-coherent baseband pulse as a video phase marker. The peaks within the pulse are centered on the appropriate group range on the ionogram display at each frequency, but the position of the baseband peaks within the pulse are controlled by the phase, producing telltale phase ionosonde "fringes." The fringes come about because the phase, and hence the position of the peaks on the video, systematically change with frequency, due to the physics of the propagation. This process can be simulated digitally by using the phase path as generated by the ray tracing, and using it to arrange peaks in a simulated pulse, centered on the ray-group range.

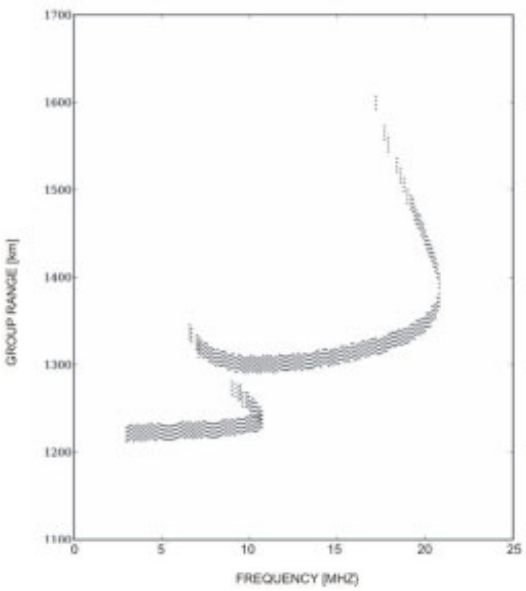

Figure 4. A simulation of a phase ionogram, O mode only, over a 1200 km path, using an IRI ionosphere. The behavior of the phase fringes can be understood by the application of Equation (5), which relates the change of phase with frequency to the group range.

A significant issue in the phase ionogram synthesis is in insuring that the ray tracing is accurate enough such that estimates of phase path are not obviously affected by accumulated error through the integration. This demands not only very accurate homing, as mentioned earlier, but very high tolerances set on the Runge-Kutta-Fehlberg differential solver. Unlike a group-path solution, where a few meters of accumulated error over the entire ray path may be inconsequential, this would have a significant impact on a technique that is trying to access the relative phase accurately within the 2π (wavelength ~10 m) ambiguity, as phase-ionosonde modeling does. Error contributions of the order of 1 mm may seem tiny, but it has been found that errors of the order of 10 cm ($\sim \lambda/100$) become visually apparent in phase ionograms, so a 1 mm error insures a negligible contribution from the homing scheme. The accumulated error along the ray path must also be similarly taken into account and constrained.

Figure 4 shows a synthetic phase ionogram, generated by this process over a 1200 km mid-latitude path through the IRI. Only the O mode is shown for clarity. The classic phase fringes of the phase ionosonde are apparent. The significance of the fringes can be explained with a little theory. For example, Budden [12] shows that the group range, h', and the phase range, p, for an ionospheric ray with center angular frequency ω, may be expressed as

$$h' = \frac{d(\omega p)}{d(\omega)} = p + \omega \frac{d(p)}{d(\omega)} . \qquad (4)$$

Using the well-known relationship between wave phase, ϕ, and phase path on a two-way path, we can find that

$$h' = \frac{c}{4\pi}\frac{d\varphi}{df}, \qquad (5a)$$

or

$$t' = \frac{1}{4\pi}\frac{d\varphi}{df}, \qquad (5b)$$

where t' is the group delay [20]. This suggests that the slope of the fringes are related to a precise measurement of group height. The relationship is not absolute, and is only unique within a 2π ambiguity in phase path. The rate at which the fringes go through repetition in pattern with frequency is itself an indicator of the rate of change of the group range with frequency. Inspection of Figure 4 illustrates this point, and there are several regions of interest within the ionogram. Note that in the E region, the repetition of phase patterns is slower than in the F region, and that the rate of change increases as the nose of the ionogram is approached. Within the F region, note that the appearance of the pattern is controlled by the regions where group range is decreasing, increasing, and stationary with frequency.

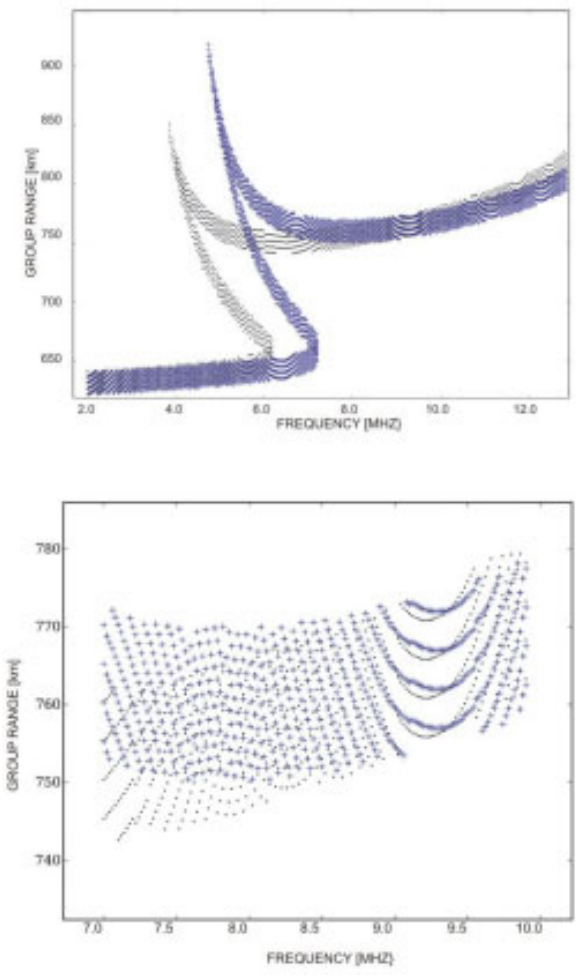

Figure 5. Two looks at a model phase ionogram with the O and X modes plotted. The lower plot (Figure 5b) has the frequency scale expanded to look at the region of the F trace where the O and X modes converge. The model has also been executed with high frequency resolution to more closely highlight differences between the two modes.

In more-complicated ionospheres – for example, where the ionosphere could not be thought of as a smooth refracting medium, due to plasma instabilities or other mechanisms – the fringe patterns of ionograms will become disturbed. The synthesis of phase ionograms in the above manner might allow models of the propagation medium to be evaluated against measured results.

Figures 5a and 5b show both ordinary (O) and extraordinary (X) modes for part of a phase ionogram over a 600 km path. The model was done with high-frequency resolution, to highlight the differences for O and X. Note that the phase fringes converged in Figure 5b when the modes came together in group range, but they were not an exact match. This illustrated slightly different group ranges for O and X at nearly all frequencies, hinting at the ability to use the phase-frequency relation to discriminate between the two, as is attempted in the next application (Section 3).

According to Equation (5), the phase ionosonde technique provides the means to get very accurate estimates of the group range by estimating the rate of change of phase. This has several potential OTHR applications. One such application is explored in the next section, where the relationship between phase and frequency might be exploited to produce a wideband waveform that can resolve direct and Earth-bounce ray paths to targets, which when combined can provide estimates of the target's altitude.

3. Application Two: Resolving OTHR Micro-Multipath

The ideas in the previous section may be extended to look at providing a phase-measurement extension to OTHR reception through the use of a wideband waveform. Most conventional OTHR uses FMCW, but the bandwidth is often limited by the need to avoid interference with other users. Schemes to extend the bandwidth and to notch interfering in-band channels result in sidelobe issues in the ambiguity function, unless specially treated [23, 24]. A waveform proposed here is the sum of a series of frequency-offset linear-frequency-modulated continual-wave (FMCW) waveforms, viz.,

$$W(t) = \sum_i FMCW_i(t)\, w(t) \cdot \qquad (6)$$

The generation process would be done by digitally constructing $W(t)$, and then programming a direct digital arbitrary waveform generator with the appropriate phase progression. The $w(t)$ are weighting functions for sidelobe control. Except for their center frequencies, the waveform parameters should be the same for each $FMCW_i$. In practice,

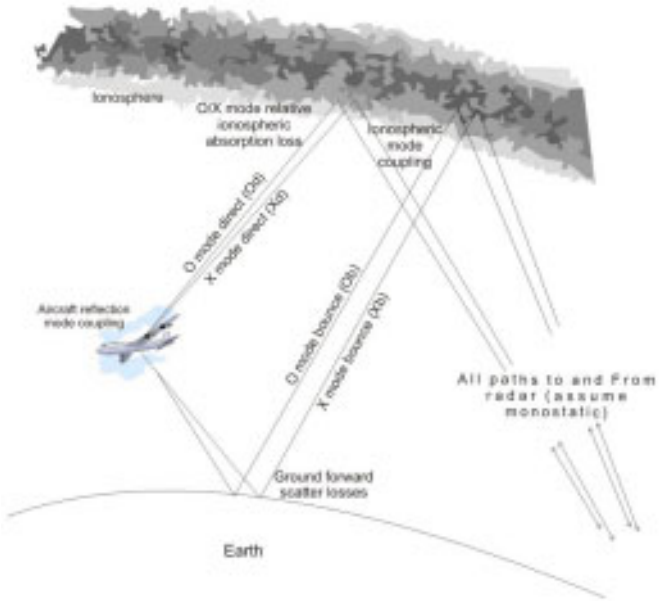

Figure 6. A schematic diagram showing the four independent one-way propagation paths from an ionospheric layer to an OTHR target. Each mode then has four potential return paths, giving 16 two-way paths. Complex physics, including forward-scatter loss, ionospheric absorption loss, and mode coupling will determine the paths' relative amplitudes.

the choice of the center frequencies and bandwidth of the $FMCW_i$ would be achieved by the use of a clear channel watcher. It is important to note that each individual $FMCW_i$ will only carry $1/i$ of the total effective radiated power of $W(t)$.

At the receiving site, each FMCW waveform is treated independently, using digital reception tuned to the various center frequencies. The received $FMCW_i$ would then be treated as usual in the processing chain. At any independent center frequency *i*, the output will be very similar to a simple FMCW, except reduced in power by $1/i$. In the case of a linear-array OTHR, the final output normally consists of spectral estimates in the dimensions of range (*r*), Doppler (*D*), and azimuth of arrival (θ):

$$P = P(r, D, \theta) \cdot \tag{7}$$

In the case of the wideband waveform, we now have

$$P = P(r, D, \theta, f) \cdot \tag{8}$$

The dimension in *f* forms a frequency series, the Fourier transform of which can be used to estimate the power estimates in a micro-range domain [22]. This can be written as

$$P \xrightarrow{F} S = S(r, D, \theta, h'), \tag{9}$$

where h' is the micro-range dimension, with its resolution now given by the overall bandwidth of $W(t)$, achieving the wideband capability.

In practice, there would be several impediments to implementing such a scheme for OTHR. In particular, there are the issues of nonlinear mixing in high-power transmitters, and the handling of interferers within a wideband waveform. The latter will lead to the need to be able to select narrowband waveforms at arbitrary center frequencies within the wideband signal. This is a principal driver for choosing the nature of the waveform and the method described here, which could readily accommodate this. The transmitter issue would require assessment for a practical implementation. However, this paper is more concerned with the initial step of demonstrating the utility of the concept and its natural limitations as imposed by the ionosphere and propagation, and therefore will not concern itself with this matter. To progress, a model is developed to explore the concept.

Consider an aircraft flying at a specified height in the footprint of an OTHR operating in the proposed micro-ranging mode, with $W(t)$ centered on a band that is, in an approximate sense, optimized for first-hop-propagation sensitivity (see Figure 6). Let us assume that the ionospheric layer providing the propagation is relatively smooth, with variations in the electron concentration that are small over the scale of a Fresnel zone, i.e., a medium that appears smooth to an HF transmission.

This single, smooth layer of the ionosphere may still provide numerous modes of energy propagation to the target. The ionosphere is a magneto-plasma, which confines propagating radiowaves to so-called characteristic modes. These modes are defined by the relative geometry of the wave normal and the geomagnetic field, as well as the concentrations of electrons, ions, and neutral densities (e.g., [12]). OTHR antenna design ignores the polarization associated with these modes, and transmits a linearly polarized wave, due to geometric convenience at the ground.

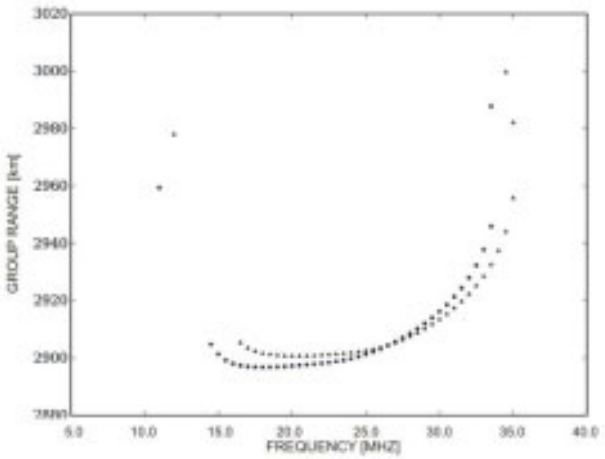

Figure 7. An oblique ionogram over a 2800 km path. The micro-ranging technique was applied, with a 640 kHz bandwidth near 27 MHz.

The result is that on entering the ionosphere, the transmitted energy couples into both ordinary (O) and extraordinary (X) characteristic modes. These modes travel independently, have different polarizations, and will in general have different phase and group paths over a circuit.

The two-way propagation for a target at altitude will be further complicated by the fact that both O and X modes will find a direct path and a reflected path from the Earth's surface ([25, 26], and see Figure 6). With mixed paths on outward and inward legs, this will result in up to eight total separate two-way paths for the two ionospheric modes. If one allows mode coupling in the region of the target, to further generalize the model, then 16 total modes are possible to and from the target region. In terms of group path, it is irrelevant where the coupling occurs in the target region, since both modes travel with similar velocities there. Mode-to-mode coupling in the ionosphere is ignored for this model, and this is likely a good assumption for most OTHR operations. Several of the 16 two-way paths are degenerate in group range for a monostatic setup, resulting in 10 "unique" paths in group range.

With a conventional resolution of 15 kHz, most of these modes would be unresolved, under normal operating and environmental circumstances. Attempts to use the phase and amplitude variation of unresolved target signatures in OTHR for target-height estimation have largely been challenged by the lack of a clear representation of the behavior of the O and X modes. To model this scenario, the *PYTHON* program, which was used to synthesize oblique ionograms, was readily extended to provide intercept at a point above the Earth's surface, as would happen with OTHR target detection of an aircraft.

As the elevation launch angles of the direct path to the target and the Earth-bounce path are different, the homing method must be repeated separately for each path. Similarly, X and O ray tracing requires separate calculations, resulting in four runs by five rays, giving a total of 20 rays per frequency. Normally, this would be a trivial number of rays. However, when the tolerances on ray tracing have been tightened to provide the 1 mm accuracy criterion, it ends up taking about 15 minutes of a single core's computing time on a Pentium M machine per frequency. This is adequate for the modeling exercise.

Once the ray states for the four one-way paths are calculated, the two-way paths can be created. No serious attempt has been made to realistically estimate the relative amplitude of the modes at this stage. Nevertheless, the relative amplitude is likely to be an important aspect in the success of a height-finding technique using this concept, and so several physical processes were treated as input variables to the model, viz.:

- X-mode-to-O-mode ionospheric absorption ratio (dB)
- Mode coupling loss at target (dB) (i.e., the loss attributed to any mode that travels to and from the target on different modes)
- Ground forward-scatter loss (dB)

The oblique ionogram is generated as a useful first step in understanding the propagation characteristics to the target region. It may also be used as a guide to sensible frequency selection for position of the wideband (but much narrower than the oblique ionogram radar signal) transmission.

The oblique ionogram generated on a 2800 km path looking south along a magnetic meridian is shown in Figure 7. As a rule of thumb, a frequency around 80%-90% might be selected on the basis of optimizing the focusing gain on this circuit. On this basis, we chose a starting frequency of 27 MHz for this case. This frequency also put some stress on the micro-ranging technique, because the O and X modes were very close in range at this point.

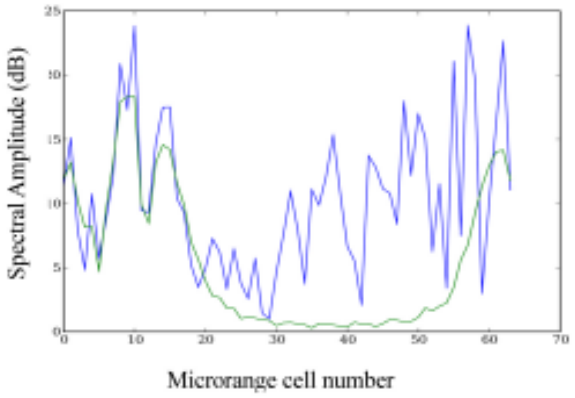

Figure 8. The micro-range spectrum, comprising 64 cells of 234 m covering 15 km.. Ten paths with unique group range provided energy to the full spectrum (blue, jagged curve), including the dominant Od-Od mode (green, lower, smoother curve). These component energies have been smeared through ionospheric dispersive effects.

The model was run with 64 adjacent channels, each with 10 kHz bandwidth, giving a 640 kHz overall bandwidth. This resulted in a nominal resolution of 234 m in the micro-ranging dimension. Figure 8 shows the spectral output using a simple un-weighted FFT. The blue curve shows the spectrum of the combined signal, and the (lower, smoother) green curve shows the spectrum of the O-mode direct two-way path, which should have been the strongest in the combined signal.

The result was promising, in the sense that much of the energy for this mode was confined to one region of the spectrum, and other components were clearly visible. However, there was also a significant dispersive effect, causing the mode to smear energy over a large portion of the micro-range space, significantly degrading the effectiveness of the wideband signal.

Future work is required to look at means to remove the dispersive effects. Inspection of Figure 7 suggests that over bandwidths of this order, the dispersion is approximately linear. This provides some confidence that basic time-frequency analysis would be effective in removing the dispersive effect, and in allowing the resolution of micro-multipath using this technique.

4. Summary

Two applications involving the phase path calculated by ray tracing based on Haselgrove's formulations have been demonstrated. The first was the simulation of oblique phase ionograms. The second was the construction of a wideband method for micro-multipath resolution in OTHR. Both techniques showed the promise of providing insight into complex propagation phenomena affecting OTHR. The wideband method as presented does not adequately deal with ionospheric dispersion, but the linear nature of the dispersion over representative bandwidths suggests that time-frequency analysis would be successful in progressing this method.

Whilst these techniques show promise, perhaps the more profound conclusion of this paper is that the legacy left by Jennifer Haselgrove through her radio ray-tracing formulation, and the subsequent developments in propagation modeling that her work inspired, are still providing areas of exploration some 50 years on.

5. References

1. J. Haselgrove, "Ray Theory and a New Method for Ray Tracing," in *The Physics of the Ionosphere: Report of the Physical Society Conference on the Physics of the Ionosphere, held at Cavendish Laboratory, Cambridge, September 1954*, London, Physical Society, 1955, pp. 355-364.
2. J. Haselgrove, "Oblique Ray Paths in the Ionosphere," *Proc. Phys. Soc.*, **70B**, 1957, pp. 653-62.
3. J. Haselgrove, "The Hamiltonian Ray Path Equations," *J. Atmos. Terr. Phys.*, **25**, 1963, pp. 397-9.
4. C. B. Haselgrove and J. Haselgrove, "Twisted Ray Paths in the Ionosphere," *Proc. Phys. Soc.*, **75**, 1960, 357-63.
5. J. M. Kelso, *Radio Ray Tracing in the Ionosphere*, New York, McGraw-Hill, 1964.
6. C. J. Coleman, "A Coordinate Registration Algorithm Based on Numerical Ray Tracing," Research Report RR-0009, DSTO, Edinburgh, Australia, 1994.
7. R. I. Barnes, "Real Time Ionospheric Models for the Australian Defence Force," WARS2000, http://www.ips.gov.au/IPSHosted/NCRS/wars/wars2000/commg/barnes.pdf , 2000.
8. R. I. Barnes (ed.), *Proceedings of the 1999 US-AUS Memorandum of Agreement on Radar CR Workshop*, DSTO-DDP-0477, 625 pages, 1999.
9. L. J. Nickish, S. V. Fridman, and M. A. Hausman, "Automated Propagation Assessment for OTHR Final Technical Report," Tech Rep MRC/MRY-R-070, Mission Research Corporation, 1994.
10. C. J. Coleman, "A Ray Tracing Formulation and its Application to Some Problems in Over-the-Horizon Radar," *Radio Science*, **33**, 4, 1998.
11. R. I. Barnes, "Faraday Rotation in a Cold, Inhomogeneous Magneto-Plasma: A Numerical Comparison of Ray and Full Wave Analyses," *Radio Science*, **32**, 1997, pp. 1523-1532.
12. K. G. Budden, *The propagation of Radio Waves*, Cambridge, Cambridge University Press, 1985.
13. I. N. Capon, "The Application of Ray Tracing Methods to Radio Signals from Satellites," *Proc. Phys. Soc. London*, **77**, 1961, pp. 337-45.
14. M. G. Golley, "The Tracing of Radio Waves in the Ionosphere," Technical Note Pad 96, WRE Australia, 1965.
15. R. M. Jones, "A 3-D Ray Tracing Computer Program," *Radio Science*, **3**, 1968, pp. 93-4.
16. J. A. Bennett, P. L. Dyson, and R. J. Norman, "Progress in Radio Ray Tracing in the Ionosphere," *The Radio Science Bulletin*, No. 310, September 2004, p. 81.
17. D. Bilitza, "International Reference Ionosphere 2000," *Radio Science*, **36**, 2, 2001, pp. 261-275.
18. J. H. Mathews, *Numerical Methods*, London, Prentice-Hall International, 1987.
19. H. J. Strangeways and R. T. Ioannides, "Ionospheric Effects on Earth-Satellite Paths Using Mqp Modelling and Nelder-Mead Optimisation," *Proceedings IEE National Conference on Antennas and Propagation*, 1999, pp. 196-199.
20. J. D. Whitehead and A. Malek, "A Suggested Method of Accurately Measuring the Virtual Height of Reflection of Radio Waves from the Ionosphere," *J. Atmos. Terr. Phys.*, **25**, 1961, pp. 599-601.
21. J. D. Whitehead and E. Kantarizis, "Errors in the Measurement of Virtual Height of Reflection of Radio Waves in the Ionosphere," *J. Atmos. Terr. Phys*., **29**, 1967, pp. 1483-8.
22. J. C. Devlin, P. L. Dyson, P. R. Hammer, "A Pulse Synthesis Technique for Improving Ionosonde Resolution," *Radio Science*, **12**, 5, 1977, pp. 767-72.
23. D. Zhang and X. Liu, "**Range Sidelobes Suppression for Wideband Randomly Discontinuous Spectra OTH-HF Radar Signal,"** ***Proceedings of the 2004 IEEE*** *Radar Conference*, April 26-29, 2004, pp. 577-581.
24. D. Zhang and X. Liu, "A Sidelobes Suppression Technique for Spectra Discontinuous HF Radar Signal Based on Spectra Compensation Algorithm," *Proceedings of the 2004 IEEE Asia-Pacific Radio Science Conference 2004*, pp. 242-245.
25. M. Papazoglou, J. Krolik, "Electromagnetic Matched-Field Processing for Target Height Finding with Over-the-Horizon Radar," *1997 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'97)*, **1**, p. 559.
26. M. Papazoglou and J. L. Krolik, "Matched-Field Estimation of Aircraft Altitude from Multiple Over-the-Horizon Radar Revisits**,"** *IEEE Transactions on Signal Processing*, **47**, 4, 1999, pp. 966-976.

GPS : A Powerfull Tool for Time Transfer

P. Banerjee

Abstract

The use of the Global Positioning System (GPS) for navigation and timing has been increasing with an accelerating pace since the system's inception. This has been catalyzed by the availability of different types of GPS receivers, suitable for various applications, and by a sustained effort to improve the accuracy capability of the system. Thus, in the course of time, GPS has become popular for timing applications, as well. This paper elaborates the different techniques that have evolved to cope with the demand for higher accuracy, and describes the respective merits and demerits. It also outlines the future strategies for GPS vis-à-vis the evolution of similar examples of a global navigation satellite system (GNSS), such as the Russian Glonass, the European Galileo, along with a few regional augmentation endeavors.

1. Introduction

The fundamental component of timekeeping is a clock. A clock consists of a frequency standard, a counter for counting oscillations, and a display device for displaying the count. A clock must also have a provision for setting the time with respect to a reference time (i.e., national standard time). The oldest frequency reference for a clock has been the rotating Earth. The rotation of the Earth around its own axis is defined as one day: more precisely, one solar day. A more recent clock makes use of the inherent characteristics of the electronic transition from one state to another of a particular atom. This led to the development of the widely used rubidium clock, hydrogen maser, and cesium clock [1]. These are now commercially available. The cesium atomic clock has been established as the primary standard for frequency and time. The Bureau International des Poids et Mesures (BIPM – International Bureau of Weights and Measures) generates International Atomic Time (TAI) by linking the standard clocks maintained at different timing laboratories across the globe. From time to time, a leap second is applied to TAI to generate Universal Coordinated Time (UTC) [2]. This is steered to match the time observed by the rotation of the Earth about its own axis.

The last century has witnessed a rapid growth of time-transfer techniques. Consequently, many new schemes have been proposed, and some implemented, for transferring time and frequency from one geographic location to another. These schemes vary from the physical transportation of precision clocks, to the utilization of electromagnetic emissions from ground-based as well as Earth-satellite sources. For economic reasons, most of the latter schemes involve piggybacking of the time services onto existing or proposed communications, navigation, or other systems. Time and frequency signal-dissemination techniques [3] operate through carrier frequencies below 30 MHz to be reflected by the ionosphere, with those above 30 MHz being sufficiently high in frequency to penetrate the ionosphere. Signals from the former include primarily HF broadcasting and LORAN-C. These signals may be observed at great distances from the transmitter, but they suffer from ionospheric-propagation anomalies that limit the accuracy. TV and satellite signals are restricted to line-of-sight applications, but they show little or no signal deterioration caused by propagation anomalies. The most accurate systems are those that use the higher line-of-sight frequencies.

In the HF band, electromagnetic signals from 3 MHz to 30 MHz have been used. Standard time signals, broadcast via HF, have long been a favorite tool for time transfer since the early 1900s. During times of good radio propagation, any simple radio with a shortwave (SW) band can be used to pick up standard time and frequency signals (e.g., ATA, JJY, WWV, etc.) [4]. These usually consist of an audio tone, a burst of an audio signal every second, a voice announcement at some regular interval of time, etc. Uncertainty in determining the propagation delay generally restricts accuracy to the order of one millisecond, the best accuracy that can be relied upon for time markers transmitted on HF carriers along skywave paths. Within the accuracy limitations, frequency calibration and clock synchronization can be achieved quite conveniently through HF standard broadcasts, using relatively simple and inexpensive equipment at the receiver end. By carefully choosing the frequency, the mode of propagation, and the time of day for a measurement or comparison, an observer can obtain optimum results with HF time standards. Under ideal

P. Banerjee is with the Time and Frequency Section, National Physical Laboratory, New Delhi 110012, India; e-mail: pbanerjee@mail.nplindia.org.

conditions, the attainable accuracy may be ± 1 ms. Time broadcast via HF has been very popular and has served a large number of users. However, since last decade, this has gradually lost its importance with the availability and the popularity of time services via satellites. Many countries have even phased out the respective time broadcasts via HF.

The navigational method provided by LOng RAnge Navigation (LORAN) [5] is based on the principle of the time difference between the receipt of signals from a pair of radio transmitters. A LORAN network consists of a master (M) station and, at a minimum, two slave stations (X, W). The LORAN-C navigation system was conceived as, and primarily serves as, a long-range precision hyperbolic navigation system. It typically offers users about 150 m position accuracy at ranges in excess of 1800 Km. LORAN-C chains are timed by synchronizing the transmission of the master station to the national master clock. Slave stations remain synchronized to the master station to fulfill the navigation requirement. Thus, this signal serves the purpose of time synchronization. The accuracy of the groundwave system is limited by propagation affects to about ± 1 μs. The accuracy of skywave synchronization is limited to about ± 50 μs for both daytime and nighttime.

Television services that rapidly expanded in the early 1980s offered various possibilities for time comparison and dissemination. Two basic ways of using TV [6-8] can be distinguished: the passive technique and the active technique. TV techniques make judicious use of the fact that video signals in TV during the blanking interval have synchronization pulses. These do not have a picture signal sitting above them, but they have a fixed pattern, independent of the varying content of the image. Television signals are controlled by relatively stable frequency standards to ensure a sufficiently good line and picture synchronization. This makes television signals also suitable for use in the comparison of remote clocks. This technique has been popular for quite some time for time transfer with an accuracy of few microseconds. Prior to 1981, only LORAN-C and TV links were used to compare clocks contributing to TAI.

With the availability of satellite technology, the advancement of time-transfer systems in the last two decades has been remarkable. For example, precise (100 ìs to 1 ìs) time transfer via geostationary satellite has been achieved with wide coverage. The time and frequency signal dissemination via the INSAT satellite in India, and GOES in the USA, has been operational for the last two decades [9, 10]. However, the advent of the Global Positioning System (GPS) has had a revolutionary impact in the field of navigation and timing. The advancement in technology of the GPS receiver, and the extensive research activities in improving the performance of GPS, have overshadowed all existing techniques of time transfer. Although this paper does not purport to present an exhaustive review, it attempts to outline the capabilities of GPS time transfer [11-14] by various techniques, and to explore the future potential and plans for these capabilities.

2. The Concept of GPS Time

The GPS constellation, operated by the US, has 24 satellites in three orbital planes. Each is inclined to the equatorial plane by 55°, and the planes are offset from each other by 120° in longitude. Eight satellites are in a circular prograde 12-sidereal-hour orbit in each orbital plane. With the full constellation in operation, from four to 12 satellites are always visible everywhere. At the heart of the GPS system are more than one precise atomic clocks onboard each satellite. Each satellite broadcasts spread-spectrum-coded information at 1575.42 MHz (L1) and 1227.6 MHz (L2). The L1 transmissions include a publicly available coarse-acquisition (CA) code. L1 and L2 carriers are superimposed with navigation data, including satellite ephemeris data, an atmospheric and propagation message, and satellite-clock bias information. The times of the clocks in each GPS satellite are kept closely synchronized to the UTC time scale, maintained by the US Naval Observatory (USNO) in Washington, DC. However, it may be noted that no leap second is added to GPS time. GPS time was set to match UTC in 1980, but has since diverged. The lack of corrections means that GPS time remains at a constant offset (19 seconds) with TAI. Periodic corrections are performed on the onboard clocks to keep them synchronized with ground clocks. The GPS navigation message includes the difference between GPS time and UTC, which, as of 2007, was 14 seconds. Receivers subtract this offset from GPS time to calculate UTC.

Unlike the year, month, and day format of the Gregorian calendar, the GPS date is expressed as a week number and a second-of-week number. The week number is transmitted as a ten-bit field in the navigation messages, and so it becomes zero again every 1,024 weeks (19.6 years). GPS week zero started at 00:00:00 UTC on January 6, 1980, and the week number became zero again for the first time at 23:59:47 UTC on August 21, 1999. To address this concern, the modernized GPS navigation messages use a 13-bit field, which only repeats every 8,192 weeks (157 years), and will not return to zero until near the year 2137.

While the satellites are in the orbit, there are constant movements and height changes relative to the Earth-centered inertial reference frame. According to the special and the general theories of relativity, the clocks on the satellites, are thus affected by their speed, as well as by their gravitational potential, respectively. The fractional frequency offset ($\Delta f / f_0$) in GPS clocks caused by these effects is governed by the relation

$$\frac{\Delta f}{f_0} = -\frac{1}{2}\left(\frac{v}{c}\right)^2 + \frac{\Delta U}{c^2},$$

where ΔU is the change in the gravitational potential at the satellite, v is the velocity of the satellite, and c is the velocity of light. Thus, the atomic clocks at the GPS orbital

altitude of ~26562 km will tick more rapidly, because they are in a weaker gravitational field than are atomic clocks on the Earth's surface. Atomic clocks moving at the orbital speeds of GPS satellites ($v/c = 2.7\times10^{-5}$) will tick more slowly than stationary ground clocks. These combined effects lead to a frequency offset of 4.46×10^{-10} . The frequency of the base oscillator in a GPS satellite is 10.23 MHz, from which all other frequencies are derived. Thus, to account for the relativity effect, the frequency of the base standard onboard each satellite is set at 10.22999999543 MHz, instead of 10.23 MHz, prior to launch.

In normal usage, GPS is used to find the instantaneous position of the antenna of the GPS receiver. Because GPS "time markers" serve as excellent sources for dissemination of UTC (USNO), over the last two decades GPS has also evolved into the primary system for worldwide distribution of precise time with very high accuracy, approaching a few nanoseconds (ns).

The basic principle of GPS is based on measuring the range of the satellites at the receiving ends. The pseudo range (P_i) of the *i*th GPS Satellite is given as

$$P_i = \sqrt{(x_{si} - x_r)^2 + (y_{si} - y_r)^2 + (z_{si} - z_r)^2}$$

$$+ (\Delta tu - \Delta ts)c + \Delta tai\, c + noise\,, \qquad (1)$$

where Δtu is the user clock error with respect to GPS time; Δts_i is the *i*th satellite clock error with respect to GPS time, computed from the clock-error parameters encoded in the received satellite signals; Δtai represents the propagation delays (troposphere/ionosphere), plus other errors and delays such as ephemeris errors and multipath, for the *i*th satellite (discussed in detail in Section 3); *noise* is the pseudo-range measurement noise.

The clocks of all GPS satellites are coordinated by the clock of UTC (USNO), so it may be assumed that all clocks of GPS satellites are synchronized or the difference, if any, is known to users. Thus, Δtai may be assumed to be independent of the satellite. In Equation (1), P_i is a measured quantity at the receiver.

The GPS signal data carries the orbital parameters of the corresponding satellites. From these parameters, the instantaneous position coordinate (x_{si}, y_{si}, z_{si}) of the *i*th satellite may be predicted. Thus, it turns out that in Equation (1), four parameters – the three position coordinates (x_r, y_r, z_r) of the receiving antenna and Δtu – are unknown. The basic purpose of the GPS service is to provide the instantaneous values of the position coordinates of the user and the time offset of the receiver clock. It is quite obvious that to determine four parameters, one needs to have four simultaneous measurements from four different GPS satellites. Four measurements generate four sets of equations as in Equation (1). So, for normal applications (i.e., both for position and timing), one needs to simultaneously receive signals from a minimum of four GPS satellites. Most common GPS receivers work on this principle, and give both the position coordinates and the time. For the convenience of the description, we will call this process Method I (Figure 1a).

There exists another technique of using GPS for timing applications. It is of interest to note that if the receiver position is known in advance, then one may solve for Δt (i.e., time) in Equation (1) just from the measurement from a single satellite, only one parameter being unknown. Let us call this Method II (Figure 1b). Method II is normally used by time-keeping laboratories, as the antenna has to be placed at a known, fixed point. Thus, the accuracy of GPS time will be dependent on the accuracy (elaborated on in Section 5) of known position coordinates. It may be noted that initially, when the GPS constellation was not full, four GPS satellites were not always visible at a particular place. However, timing applications were operational right from the initial launches of few satellites, since as in Method II,

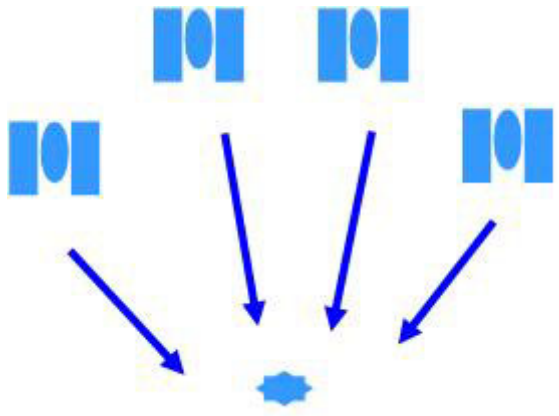

Figure 1a. In Method I, a minimum of four satellites are to be received to determine four unknown parameters (three position coordinates and time).

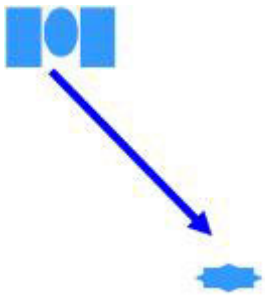

Figure 1b. In Method II, only one satellite is to be received to determine one unknown parameter, time, the three position coordinates being known in advance

one may get "time" even when only one satellite is visible. Thus, the special timing receiver with single satellite was developed right from the inception of the GPS system. Over the course of time, the single-satellite timing receiver has gone through many phases of development, both in terms of hardware technology and software improvements. This special type of timing receiver is currently widely used by time-keeping laboratories contributing to the generation of UTC, coordinated by BIPM.

Method I is applicable for any location (i.e., it does not require prior knowledge of the position of the antenna). It is more popular because of the simplicity of the technique, and because of the availability of inexpensive receivers. Because the GPS constellation is now more than completely full, the requirement of visibility of a minimum of four satellites is no longer a limitation. Furthermore, since May 2, 2000, the selective availability of GPS has been withdrawn. (Selective availability (SA) is an intentional degradation of GPS signals to deny unauthorized users access to the full accuracy. Since March 25, 1990, it had been intermittently instituted by dithering the satellite clock, and by truncating the transmitted navigation message for ephemeris. The US government has discontinued the use of selective availability since May 01, 2000.) This has thus restored the accessibility of the full potential of GPS accuracy for civilian users. In the course of time, there has been rapid advancement in GPS receiver technologies, leading to miniaturization and a reduction of the cost of the GPS receiver. All of these aspects coupled together have extensively encouraged the use of GPS for timing applications, particularly using Method I [15-18].

3. Limitations of GPS Accuracy

To assess the limits of the accuracy of GPS time, it becomes important to identify the sources of error. It is more important to look for measures to reduce and eliminate these errors. Errors are usually presented as a statistical measure of system errors (i.e., in terms of 1σ). Six major factors [19] are normally identified as the major sources of errors. These are errors in the Ephemeris data, in the satellite clock, in ionospheric modeling, in tropospheric modeling, and errors caused by multipath and receiver noise.

3.1 Ephemeris Errors

Ephemeris errors result when the GPS ephemeris data does not produce the correct satellite coordinates. While the ephemeris data is transmitted every 30 seconds, the information itself may be up to two hours old. Data up to four hours old is considered valid for calculating positions, but may not indicate the actual positions of the satellites. The ephemeris error is estimated to lie between 1.5 m to 2.5 m.

3.2 Satellite Clock Errors

The satellite's atomic clocks experience noise and clock-drift errors. The navigation message contains corrections for these errors, and estimates of the accuracy of the atomic clock. However, these corrections and estimates are based on observations, and may not indicate the clock's current state. These problems tend to be very small, but may add up to 1.5 m of inaccuracy. These satellite-clock errors affect both the C/A- and P-code users in the same way. This effect is also independent of satellite position.

3.3 Ionospheric Errors

The GPS signals are delayed [20, 21] in proportion to the number of free electrons encountered in traversing the ionosphere. This is also (to first order) proportional to the inverse of the carrier frequency squared ($1/f^2$). The phase of the radio-frequency carrier is advanced by the same amount because of these effects. The usual technique to counter this effect is to model these delays. The effective accuracy of this modeling is about 7 m.

3.4 Tropospheric Errors

Deviation of the GPS signal from the vacuum speed of light is also caused by the troposphere. Variations in temperature, pressure, and humidity all contribute to variations in the speed of light of radiowaves. For most users and circumstances, a simple model should effectively be accurate to about 0.7 m.

3.5 Multipath Errors

Multipath errors are the errors caused by reflected signals entering the receiver and masking the real correlation peak. These effects tend to be more pronounced in a static receiver near large reflecting surfaces, where 15 m or more in ranging error can be found in extreme cases. With proper site selection and antenna selection, the net impact to a moving user should be around 1.2 m under most circumstances.

3.6 Receiver Errors

Initially, most commercial GPS receivers were sequential, in that one or two tracking channels shared the burden of locking onto four or more satellites. With modern chip technology, it is common to place three or more tracking channels on a single inexpensive chip. As the size and cost have shrunk, techniques have improved, and five- or six-channel receivers are common. Most modern receivers use a reconstructed carrier to aid the code-tracking loops. Inter-channel bias is minimized with digital sampling and all-digital designs. Nevertheless, the receiver noise contributes to the error by about 1.5 m.

Each component of the errors is assumed to be statistically uncorrelated with all other components, so they are combined by the root of the sum of the squared errors. Thus, the worldwide civilian positioning error (users equivalent range error, or UERE) for GPS is about 8 m, as shown in Table 1. These values are to be multiplied by the corresponding DOP values to get the effective error in positioning and timing, respectively.

The dilution of precision (DOP) or geometric dilution of precision (GDOP) is a GPS term used to describe the effect of the geometric strength of a satellite configuration on GPS accuracy. When visible satellites are close together in the sky, the geometry is said to be weak and the DOP value is high; when far apart (i.e., with wider angular separation), the geometry is strong and the DOP value is low. The ideal value of DOP is one.

4. The Effect of Scintillation on GPS Time

In addition to errors discussed in the last section, there are natural phenomena that affect the performance of GPS timing and positioning. One such effect is scintillation, which is a matter of concern for GPS applications. GPS signals are delayed while passing through the ionosphere, and they also fluctuate in amplitude. Large irregularities in the electron-density distribution cause severe fluctuations in the signal strength, known as scintillation. The Indian subcontinent extends from the magnetic equator, touching the tip of the peninsula, to the mid-latitude zone in the north. Thus, the whole of India lies almost within the region that is prone to scintillation. It has also been well established that the strength and frequency of scintillation is highly correlated with the 11-year cycle of the sunspot number.

Scintillation [22-24] becomes a major issue for applications of the GPS network because fluctuations present additional stresses on the GPS receiver tracking loops. They can induce cycle slips, or even a complete loss of signal. The degradation of time and position solutions due to a suboptimal geometric arrangement of observable satellites is termed dilution of precision (DOP). The loss of signal from a particular satellite, which is equivalent to the absence of that particular satellite for that location, may temporarily increase the DOP value, and therefore reduces the precision of GPS timing.

(A cycle slip can occur when tracking is interrupted due to blockage of the signals, a weak signal, or incorrect signal processing due to receiver software failure. This cycle slip will alter the integer number of cycles, although the fractional phase measurement after reacquisition of the signal will be the same as if the tracking had not been interrupted.)

To study the effect of scintillations, GPS satellite observations were made in the month of November 2001 with a single-frequency multi-channel receiver at the National Physical Laboratory, New Delhi, India (NPLI). The coordinates of the antenna position at NPLI had been precisely predetermined. The time scale of NPLI (i.e., UTC(NPLI)) is linked with BIPM. The time of observation was a peak period of the solar cycle, when scintillation events were most intense, frequent, and of longest duration. As scintillation normally occurs in the post-sunset period, the observations were started well before sunset, and continued until the early morning of the next day.

A special experiment with a usual GPS receiver was conducted, particularly to study the effect of scintillations on GPS time. The residuals of the instantaneous error values of the parameters (after removal of long-term trends) were calculated. These analyses are illustrated in Figures 2a and 2b, as a function of the time of the observations. Sudden bursts of wide fluctuations were observed. It may be noted that similar types of fluctuations of latitude and time errors occurred at the same instants of time. When scintillation was absent, the fluctuations of the residuals of the time errors remained within ±10 ns. However, in the presence of scintillation, the fluctuations became as high as 60 ns. A proportional degradation in the latitude accuracy (latitude error has been converted to meters for convenience of comparison with time error) was also observed at the same instant of time. By studying the data for the entire month, it was noted that the extent of the deterioration of accuracy of GPS time also almost proportionately increased with an increase in the intensity of scintillation.

The study was carried out at NPLI New Delhi (approximately 28.6° N latitude), where scintillation is expected to be strong. At places where scintillation is expected to be even more intense (such as around the South American anomaly, −30° latitude), the deterioration of the solutions would be more pronounced. This effect is not cancelled even in the GPS common-view mode, because the time of occurrence and the intensity of the scintillation vary unpredictably, depending on the geographical location.

Table 1. The standard error model: L1 C/A.

Source	Position Error (1σ)	Time Error (1σ)
Ephemeris	2.5 m	8.3 ns
Satellite Clock	1.5 m	5.0 ns
Ionosphere	7 m	23.3 ns
Troposphere	0.7 m	2.3 ns
Multipath	1.2 m	4.0 ns
Receiver Noise	1.5 m	5.0 ns
Total	**7.85 m**	**26.16 ns**

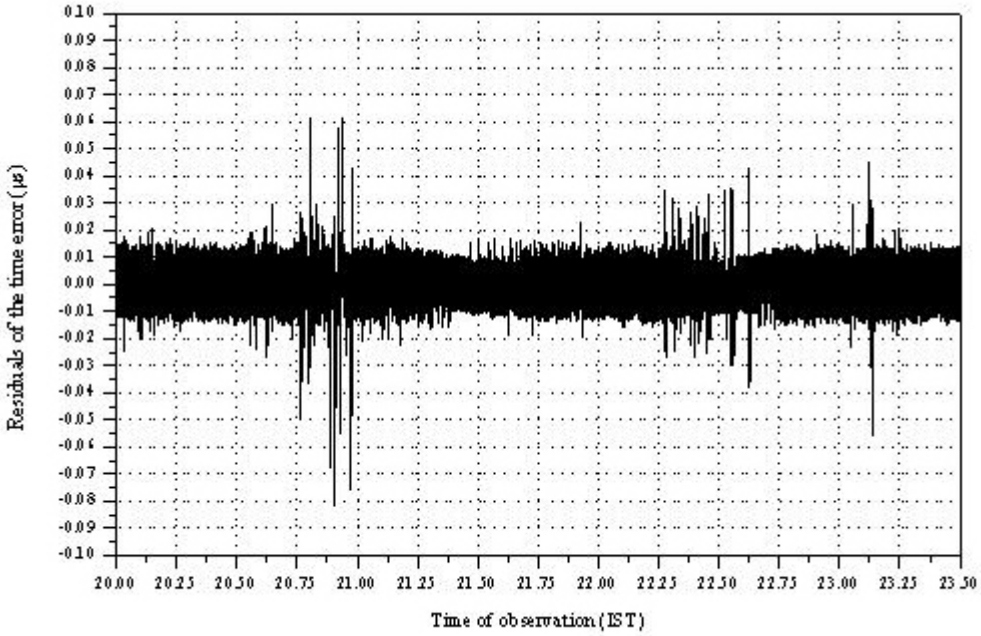

Figure 2a. The variation of the residuals of the time error (in microseconds) from a moving-point-averaged value.

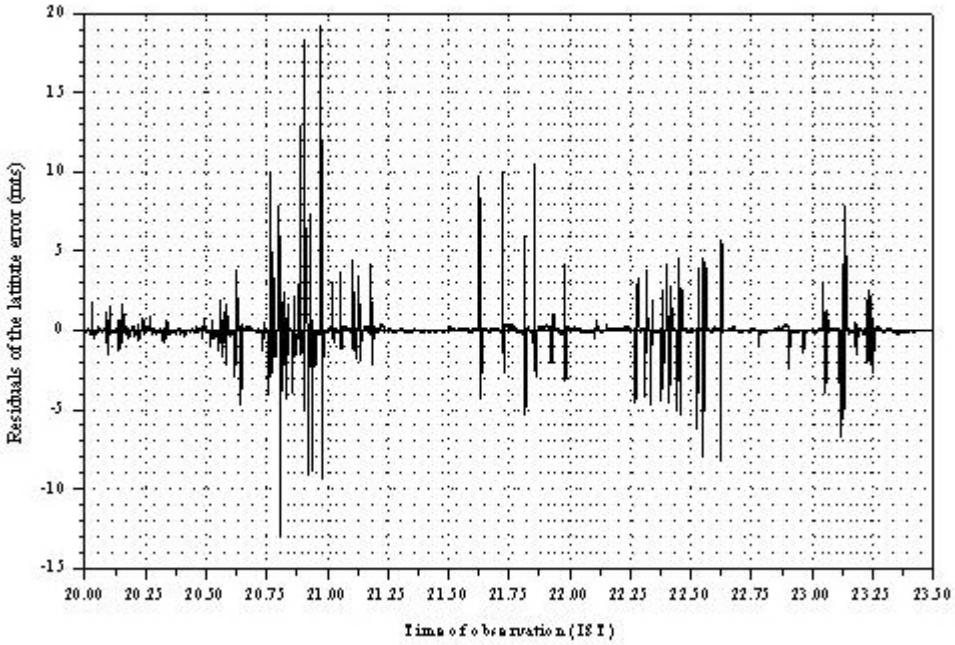

Figure 2b. The variation of the residuals of the instantaneous latitude error (in meters) from a moving-point-averaged value.

5. Effect of Position Error of Antenna in Method II

It has been pointed out that the technique of time intercomparison by Method II (Section 2) demands prior knowledge of the precise coordinates of the location. Thus, before using Method II, it is absolutely necessary to accurately determine the position of the location where the GPS receiver's antenna for the comparing clock is housed.

In some special campaign experiments at NPLI, the effect of the error in the position coordinate on GPS time was studied. The coordinates were determined very precisely in advance, within an accuracy of a few centimeters with respect to the IGS network. Two receivers were placed almost at the same locations, the coordinates of which were determined within an accuracy of 1 cm. So, the difference of time between two receivers canceled out all errors except the difference of the two receiver's noise. If deliberate errors in the position coordinates were introduced in one of

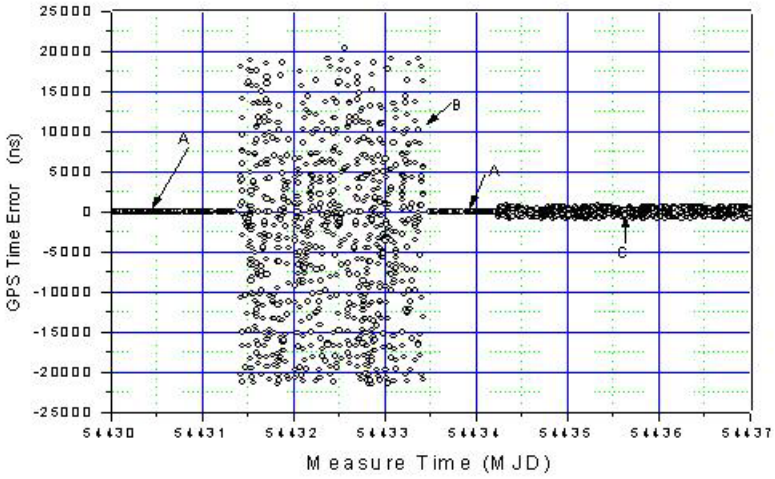

Figure 3. The effect of GPS time error on the accuracy of the position coordinate: A, for correct position (less than 5 cm); B, GPS time error exceeds even 2 ms for latitude and longitude accurate up to a minute of arc; C, errors limited to a few hundreds of nanoseconds for latitude and longitude accurate up to a second of arc.

the receivers, the difference between the times of the two receivers would thus be contributed primarily by the errors in the coordinates.

Measurements were carried out in three stages, with very precise coordinates, moderate errors in coordinates, and with large errors in coordinates. The findings of these observations (with respect to observation time in MJD) are illustrated in Figure 3. It was expected that the projection of the position-error vector on the respective satellite range vector would be directly proportional to the time error. The timing error for a particular position error would be a function of the elevation and azimuth of the corresponding satellite. It is important to note that the coordinates should necessarily be known within an accuracy of a few centimeters if one attempts to achieve a sub-nanosecond timing accuracy.

(The start of the *Julian era* is defined as the beginning at noon on Monday, January 1, of year 4713 B.C. A *modified Julian day* (MJD), which is created by subtracting 2400000.5 from a Julian-day number and thus represents the number of days elapsed since midnight (00:00) Universal Time on November 17, 1858, is usually referred to as the date of observation time in a timing experiment.)

Table 2. A Model I GPS receiver (receives only GPS).

No. of Samples (*N*)	UTC – GPS Time (ns)	1σ (ns)
5113	–136.92	51.58
5538	19.31	70.88
86847	–129.11	57.47
10105	–113.31	40.22
52670	–71.73	46.65
79908	–74.47	51.90
58278	–71.65	48.08
55872	–70.65	52.23
57363	–81.76	55.52
49815	–103.43	53.80

6. Performance of the Usual GPS Receiver for Timing: Method I

For timing purposes, and particularly by Method I, different types of GPS receivers have also evolved and are consequently also available on the market. GPS receivers available for normal timing applications may be broadly categorized into two groups. One group is the usual GPS geodetic receivers, which not only show position coordinates, but also have a hardware output of one pulse per second. The other type is a GPS-disciplined quartz/rubidium clock. The potential of Method I may be assessed by studying the performance of these receivers. To evaluate the performance of a GPS-based timing system, GPS time is normally compared with a local time, which should have a known correlation with UTC (USNO) or UTC. Studies on the performance of these receivers were carried out from time to time by a few laboratories, including the National Physical Laboratory, New Delhi, India (NPLI).

Table 3. A Model II GPS receiver (receives both GPS and GLONASS).

Mode	No. of Samples (*N*)	UTC – GPS Time (ns)	1σ (ns)
GPS	4064	1830.67	19.63
	1432	1775.33	13.17
	3818	1771.69	21.44
	63385	1733.96	37.36
	88516	1757.31	46.68
	3894	1790.77	21.36
	2675	1812.74	14.91
GLONASS	988	1216.1	22.69
	4113	1242.95	66.63
	926	1181.88	09.95
	2445	1175.13	08.13

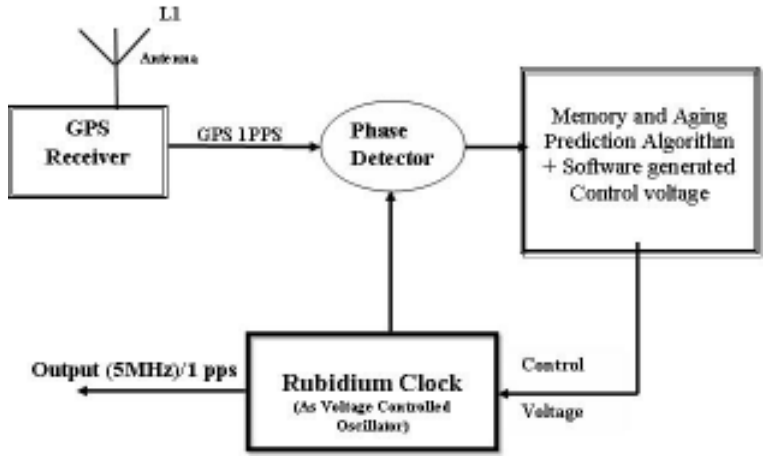

Figure 4. A functional block diagram of a GPS-disciplined rubidium clock

The performance of a clock system may be fully characterized [57] by the analysis of the values of phase-time: the fractional frequency offset (y) and the Allan deviation (σ) (defined in the Appendix). It may also be noted that these studies need well-planned observations of a sufficient number of samples of phase data (x) (see the Appendix). The performance of two distinct types of GPS receivers (one was the usual geodetic GPS receiver, and the other was a GPS-based clock) is presented separately in the following sections.

6.1 Usual GPS Receiver

With the availability of the GLONASS constellation, two types of GPS receivers are available. One receiver is the normal GPS receiver (say, Make I), and the other (say, Make II) can receive signals from both the GPS and GLONASS constellations. Tables 2 and 3 summarize the performance of these two types of receivers. A close inspection of this data revealed that the GPS time of the Make I GPS receiver varied between +19 ns to −136.92 ns. These values were well within the claimed specifications for GPS time. However, for Make II, the GPS time was more than 1 μs away from the desired value. This implies that in the design of Make II, sufficient effort was not made to adjust the delay so that it closely matched UTC (USNO).

This analysis [18] led to the observation that the normal GPS receiver needs to be calibrated with respect to any standard clock that is linked to UTC. Otherwise, the absolute accuracy of GPS time cannot be relied upon.

6.2 GPS-Disciplined Rubidium Clock

The GPS-disciplined rubidium clock [16-18] is also a GPS receiver, but it is quite different from the normal GPS receiver discussed earlier. This receiver is a combination of a rubidium clock and a GPS receiver, as shown in the functional block diagram of Figure 4. The clock is locked to the GPS time generated by the receiver. The history of the phase-comparator output that controls the rubidium clock is recorded in memory. The recorded history of the comparator's output helps in mimicking a GPS signal at least for some time while the GPS signal is not actually available. The system remains in a locked condition in two

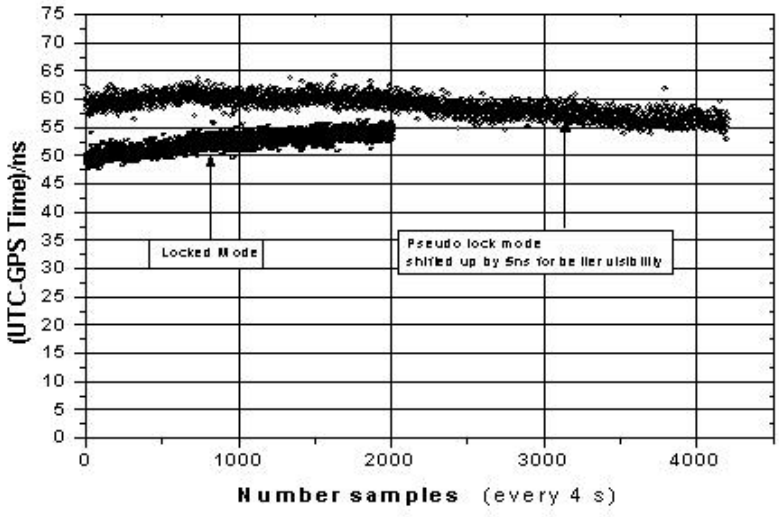

Figure 5. Phase observations in different modes of operation of the GPS-disciplined rubidium clock.

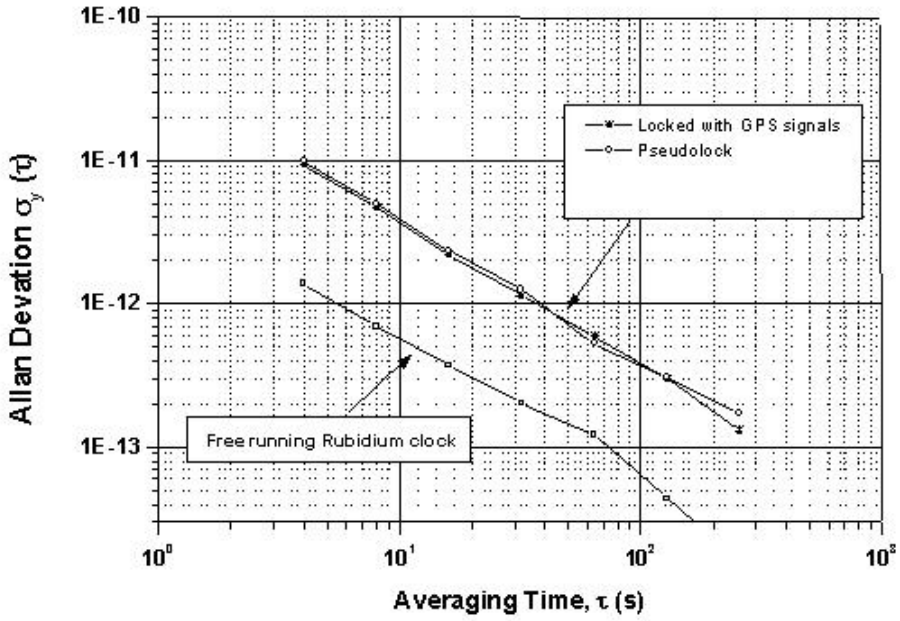

Figure 6. The frequency stability of the GPS-disciplined rubidium clock in different modes of operation.

distinct modes. In one mode (say, Mode 1), the rubidium clock achieves a lock with the GPS signals. If the GPS antenna is removed or the GPS signals are absent, the rubidium clock still maintains a "pseudo lock" for some time, based on the prior history of the recorded data. This is the other mode (say, Mode 2). A few hours after the withdrawal of the antenna, the rubidium clock also loses pseudo lock. Proper evaluations of such receivers require the study of each mode of operation. The result of one such evaluation campaign is illustrated in Figure 5. In Mode 1, the difference (UTC – GPS time) lies between 45 ns and 55 ns. This value is well within the specification of GPS capabilities, which claim a timing accuracy of better than 100 ns. The standard deviation of the residual of a linear fit during this mode was approximately 1 ns. When the GPS antenna was removed (i.e., in Mode 2), while the unit showed a pseudo lock, a steady phase was found to be maintained for almost three to four hours, after which the phase data showed a slow drift. From the time-offset data, which represents x, the frequency offset, y, can be determined with the help of Equation (5). A pair of samples of $x(t)$, taken at an interval of 400 s, was used in Equation (5) to find $y(\tau)$ over an averaging time τ of 400 s. It was seen that the frequency offset, y, was found to vary between 4×10^{-13} and 9×10^{-14} when the system was in locked mode. So, in locked mode, the frequency of the rubidium clock closely followed that of UTC. However, once the lock was lost, the frequency of the rubidium clock restored to its free-running mode, and attained a frequency offset of $\sim 8\times10^{-11}$, which was quite far off from UTC. By using the same phase data (i.e., $x(t)$), the Allan deviation, which is representative of the frequency stability, was determined using Equation (6) supplemented by Equation (7). This is shown in Figure 6.

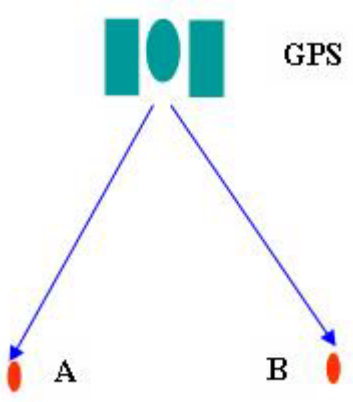

Figure 7. The concept of the GPS common-view technique

The stability of the locked clock lies between 1×10^{-11} and 1×10^{-13}, and it improves with the averaging time, τ, as seen in Figure 6. However, the stability of the free-running rubidium clock is never higher than 10^{-12}. It is interesting to note that the stability of the free-running clock is better than that in locked condition for averaging times less than 300 s. This may be attributed to the fact that in the absence of the GPS signal, the rubidium clock has less noise. In the presence of the GPS signal, the rubidium clock gets locked to the GPS signal, improving the frequency accuracy. However, the noise in the received signal deteriorates the frequency stability, particularly for a relatively shorter averaging time (less than few hours). However, as long as the rubidium clock remains locked to the GPS signal, the accuracy of the frequency remains almost same as that of the GPS timing system. The frequency of the rubidium clock will remain almost unchanged in the presence of the GPS signal, thus arresting the aging of the rubidium clock. This implies that the long-term stability (for an averaging time of more than one day) of the disciplined oscillator improves significantly over that for a free-running rubidium clock.

A GPS receiver provides a 1 pps output, which is normally correct within 60 ns with respect to UTC. However, it has also been seen that the GPS time in a GPS receiver is sometimes not well within specification. Thus, the receiver needs to be calibrated prior to precise timing applications. It has been seen that the frequency accuracy of the rubidium clock improves substantially in the presence of the GPS signal. It has been observed that the performance in the pseudo-lock condition is almost as good as that in the locked condition. This implies that the system is quite reliable for one or two hours, even after the GPS signal is absent. In particular, this characteristic may be quite useful in some typical strategic applications. It may be pointed out that for a particular receiver, the phase might have some bias error, but it does not affect the frequency, the phase bias being constant.

7. Use of a Single Satellite for Timing: Method II

Method I, as explained in Section 6, is one of the ways to get direct access to GPS time. Method II may also be used in standalone mode for direct access to GPS time. However, it requires that the receiver must be placed at a predetermined location. This makes this technique immobile. Furthermore, this also requires a special receiver, which is not very commonly in demand and is thus quite expensive. These limitations do not encourage the use of this method for direct-access applications. Thus, Method II is normally used for precise comparison of two remote clocks in a common-view mode.

7.1 Common-View (CV) GPS

The common-view (CV) method compares two remotely located clocks by viewing one common GPS satellite. The basic concept of the common-view technique is illustrated in Figure 7. A GPS satellite serves as a single common transmitter that is in common view. The common-view technique [26] has long been used for precise international comparisons of time and frequency standards.

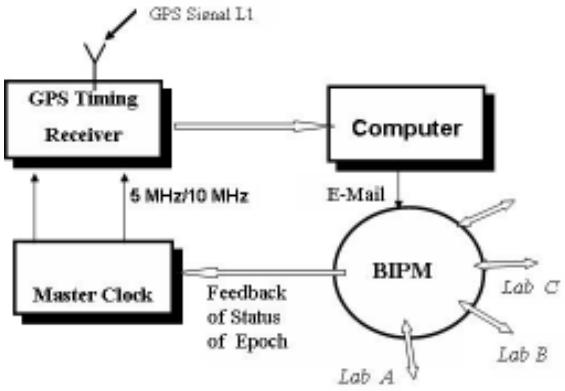

Figure 8. The concept of and International Time Link through GPS common view.

The received GPS time from each of two receivers is compared with the local clock, and the corresponding data are recorded. At location A, the compared data correspond to $(\text{GPS Time A} - \text{Clock A})$. Similarly, at location B, one records $(\text{GPS Time B} - \text{Clock B})$. The difference between these two measurements is

$$(\text{Clock A} - \text{Clock B})$$

$$- (\text{GPS Time A} - \text{GPS Time B}) \cdot$$

In common-view mode, the satellite-clock error is independent of the location of the receiver, so it is totally cancelled. Errors contributed by other factors are reduced to a varying extent, depending on the distance between the two receiving sites, called the baseline. For example, for a short baseline, the contribution by ionospheric and tropospheric delay errors may be largely reduced. So, if measurements are taken in a strict common-view mode (i.e., the same satellite is observed at the same time), the factor $(\text{GPS Time A} - \text{GPS Time B})$ is appreciably reduced for a short baseline. If the sites are separated by a long baseline, receiving conditions tend to be different at each site. The effects of the ionosphere and troposphere do not cancel out at all when the data are subtracted. In some cases, even measurement uncertainties will be equivalent to those obtained using the one-way technique. To get the best results, it is also desirable to use identical receiving equipment and antennas at both sites, and to survey the antenna positions as accurately as possible (as elaborated in Section 5).

7.2 GPS Common View for Generation of TAI and UTC

The BIPM (International Bureau of Weights and Measures) computes TAI (International Atomic Time) and UTC by using inter-comparison data from atomic clocks kept at laboratories/institutions scattered all over the world. Around 400 clocks from 59 laboratories participate in the evolution of TAI and UTC [62]. Most of the laboratories are linked through GPS common-view mode. In GPS common-view mode [25-28], most of the participating laboratories have the setup shown in Figure 8.

The GPS receiver is connected to the 1 pps (one pulse per second) signal delivered by the local standard clock. Following a procedure recommended by the Consultative Committee of Time and Frequency (CCTF), the internal software of the receiver computes the clock offset between $\text{UTC}(k)$ (i.e., for the kth laboratory) and GPS time as realized by each satellite. Figure 9 shows all participating clocks that contribute to the scheme coordinated by BIPM. Thus, it is necessary that measurements of a particular zone be made simultaneously with those of another zone through one common satellite.

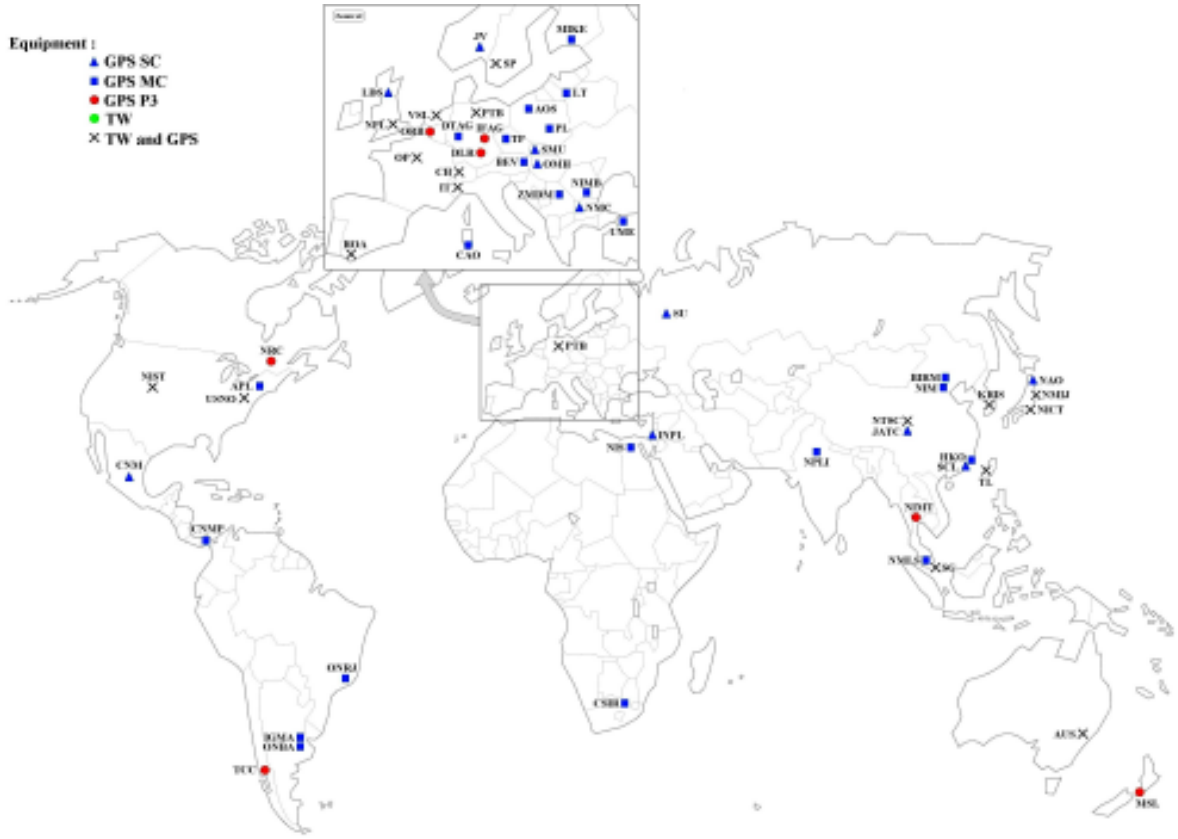

Figure 9. The participating laboratories are well scattered over the world, with a greater concentration of laboratories in Europe.

In common-view mode, it is necessary that both comparing stations follow a common tracking schedule, so that the receivers at both locations make measurements from the same satellite at the same time. One way to achieve this is with the help of a single-channel (i.e., one-satellite-at-a-time) receiver. In the early 1980s, only single-channel receivers were available. A few GPS satellites were in orbit during the early 1980s, and initial receivers were made to track only one satellite at a time. Single-channel common-view GPS receivers were first developed at NIST (now US

Table 4. GPS Schedule No. 49 (November 2, 2007).

Hawaii					Australia					India				
Class	PRN	Start		Connects	Class	PRN	Start		Connects	Class	PRN	Start		Connects
		h	m				h	m				h	m	
38	31	0	10	I	CC	19	0	42	SAF	38	31	0	10	H
D4	31	0	26	I	AC	03	1	14	I	D4	31	0	26	H
34	06	0	58	WNA,EA,ENA	AD	03	1	30	I	A0	23	0	42	ME,EA,E
35	06	1	14	WNA,EA,ENA	AE	03	1	46	I,EA	A4	20	0	58	SAF
34	07	1	30	WNA,EA,ENA	AC	19	2	2	I	AC	03	1	14	A
D4	06	2	2	WNA,ENA,	CC	11	2	18	SAF	AD	03	1	30	A
D5	06	2	18	WNA,ENA,	3C	31	2	34	H,EA	AE	03	1	46	A,EA
3C	31	2	34	A,EA	AD	19	2	50	I,EA	AC	19	2	2	A
35	07	3	6	WNA,ENA	3C	22	3	22	H	A0	13	2	18	ME,EA,E
3C	22	3	22	A	CC	20	3	38	SAF	F1	03	2	34	I
D6	06	4	10	WNA,ENA,SAM	CD	20	4	10	SAF	AD	19	2	50	A,EA
34	18	4	42	WNA,ENA	3C	16	5	30	H	A0	25	3	6	ME,EA,E
3C	16	5	30	A	CC	13	6	2	SAF	A1	25	3	22	ME,EA,E
34	22	6	18	WNA,ENA	CD	13	6	18	SAF	A0	27	3	38	ME,EA,E
38	11	7	22	I	D4	13	6	34	SAF	A1	27	3	54	ME,EA,E
3C	20	7	38	A,EA	AC	25	6	50	I	A2	27	4	10	ME,EA,E
80	03	7	54	WNA,A	3C	20	7	38	H,EA	F1	13	4	26	I
34	31	8	10	WNA,ENA,SAM	80	03	7	54	WNA,H	A0	08	4	42	ME,EA,E
35	31	8	26	WNA,ENA,SAM	AC	28	9	14	I,EA	F1	08	4	58	I
D4	31	8	42	WNA,ENA,SAM	3C	23	9	46	H,EA	A1	08	5	14	ME,EA,E
3C	23	9	46	A,EA	3C	17	11	6	H,EA	A2	08	5	30	ME,EA,E

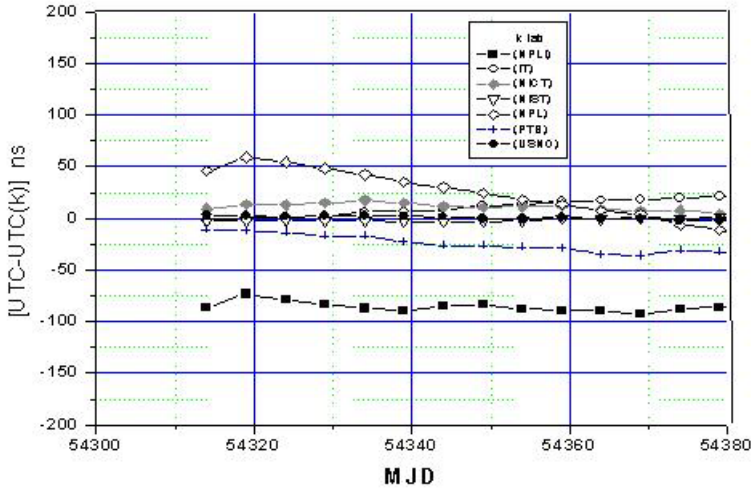

Figure 10. The recent status of the time scale of some laboratories, as given in Circular T.

National Institute of Standards and Technology; National Bureau of Standards at that time) in 1980. Several companies soon developed similar products. The single-channel technique is still widely used for international time comparisons. BIPM generates the schedule of the measurements for each zone of the world judiciously, so that the schedule of a particular time zone is always common with that of one (at least one) or more zones/regions of the world. BIPM publishes a tracking schedule for single-cannel common-view every six months.

According to the schedule, each satellite is tracked for 13 minutes. Thus, a maximum of 48 sets of measurements can be recorded per day. A typical sample of the schedule (part of the schedule No. 49) for Hawaii, Australia, and India is shown in Table 4. The fourth column of the schedule of each country indicates the zones of common view for the respective schedule. For example, for India, at 0042 hrs UTC for satellite PRN No.23, the schedule is in common view with the Middle East (ME), East Asia (EA), and Europe (E).

Now, the new generation of multi-channel timing receivers tracks many satellites simultaneously, depending on the channel capacity of the receiver (e.g., an eight-channel receiver generates eight solutions of time from eight satellites simultaneously). Over the course of time, multi-channel timing receivers [32] have been developed and are now available in the market. The multi-channel technique has the following distinct advantages:

1. For longer baselines, all visible satellites may not be in common view at both the sites, but only a few of them or only one may be common. Thus, there may not be a gap in common view, and so also in the measurement,

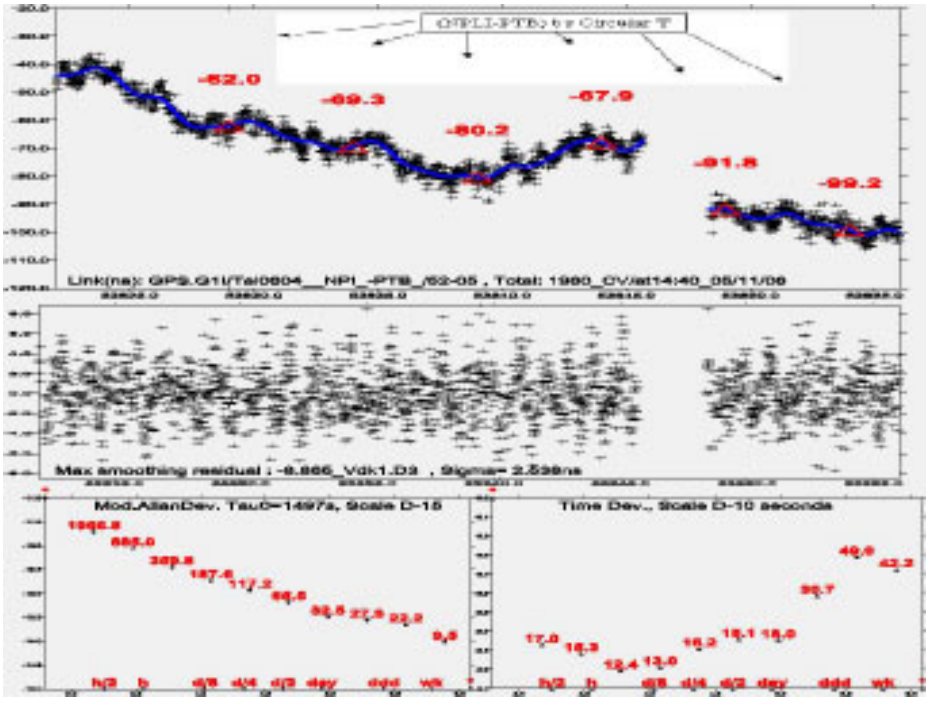

Figure 11. obtained by multi-channel GPS common view.

unlike the possibility of a gap in the common view in single-channel mode. Therefore, it is often possible for multi-channel users to continuously compare standards, with no gaps in their measurement data.

2. Because of "1," "to follow a schedule" is no longer mandatory. Thus, multi-channel common-view mode does not use a schedule.
3. In a multi-channel system, about 450 tracks can be recorded in one day using the BIPM format. This implies that in multi-channel mode, the number of tracks becomes about 10 times more than with the single-channel method. Thus, one gets more than one solution (as many solutions as the number of satellites tracked) for GPS time corresponding to each time of measurement.

However, irrespective of the type of receiver or method, it is absolutely necessary to record the measurement in one common format (also conceived by BIPM), for convenience of operation and for coordination. The clock offsets are recorded by the receiver in a particular format, known as the Common GPS GLONASS Time Transfer Standard (CGGTTS).

The status of UTC(k) with respect to UTC is published through the monthly publication "Circular T" by BIPM. The ultimate goal of each national time laboratory is to keep a local representation of UTC, UTC(k), in close agreement with UTC (say, within 100 ns). The recent status of UTC(k) for some of the laboratories is shown in Figure 10. In addition to the status of UTC(k), Circular T also contains the uncertainties (Type A and Type B) of measurements.

We discuss the link between UTC(NPLI) and the UTC system through a time link to Europe, namely to UTC(PTB) Germany, to elaborate on the uncertainty of the link. Figure 11 illustrates these analyses. After applying International GNSS Service (IGS) products of ionospheric data and precise ephemeredes, the Type A uncertainty for GPS multi-channel time transfer between NPLI and Europe was 2.5 ns. White phase noise due to the time-transfer technique was observed up to an averaging time of about 1.5 day. Thus, the data could be smoothed with a cutoff period of about 1.5 day. Initially, the Type B uncertainty of UTC (NPLI) was 20 ns. The TTS2 receiver, which was recently procured by NPLI, was calibrated at BIPM before delivery by the manufacturer. Thus, the corresponding uncertainty of Type B was reduced to 7 ns. Accordingly, per BIPM Circular T Section 6, the NPLI/PTB time link had an uncertainty of Type A of 2.5 ns, of Type B of 7.0 ns [29], and the total uncertainty was 7.4 ns. If the receiver of NPLI is calibrated at NPLI by the BIPM receiver, then the Type B uncertainty may be further reduced from 7.0 ns to 5.0 ns.

In the common-view scheme for a long-distance link, intermediate laboratories are needed to bridge the time transfer. Thus, there are four pivot laboratories (PTB, NICT of Japan, NIST and USNO of the USA) for the pivot-bridging configuration. As an example, to link PTB and AUS (Australia), it is necessary to use NICT to bridge the link.

8. All-in-View (AV) Technique

The common-view (CV) method is used in almost all the GPS time-transfer practices, since for very short baselines, dominant sources of errors are cancelled. Since March 2004, the International GNSS service (IGS) [31] has aligned its time product to GPS with the evolution of a new product, IGST, a new realization of GPS time. IGST is completely free of the broadcast GPS time. The IGS provides precise ephemeredes and clock solutions for the GPS satellites. Prior to the availability of IGS products, these were the major sources of errors, which were greatly reduced by common view. So, since March 2004, common view has lost its advantages for improved uncertainties.

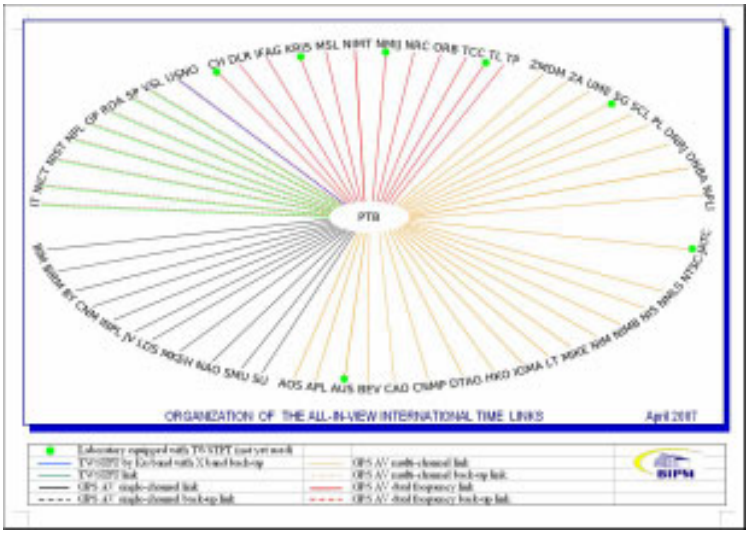

Figure 12. The current status of the international time link with PTB as the pivot laboratory.

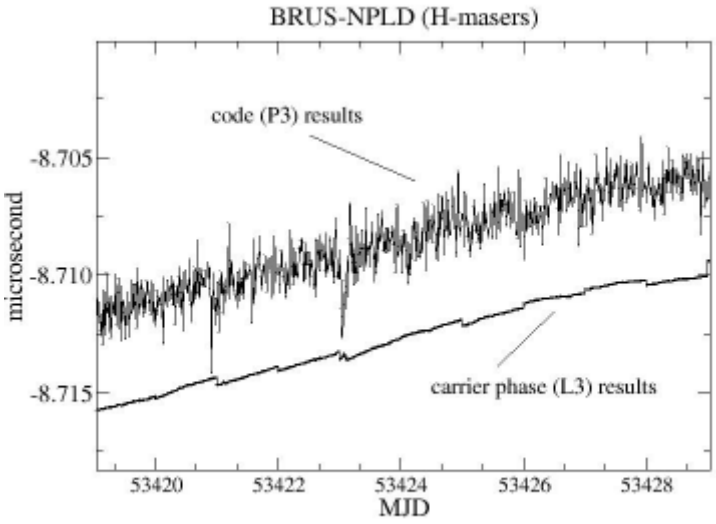

Figure 13. A comparison between time transfer with precise code and combined code-carrier phase analysis (the two curves have been separated by 5 ns in order to improve the visibility [34]).

Furthermore, the basic inherent limitations of common view have been that as the baseline increases, the number of common-view satellites decreases. With the availability of multi-channel timing receivers, another technique, called all-in-view (AV), is being studied in addition to common view for time comparison. In all-in-view, a laboratory observes all visible satellites for each 13-minute track. Thus, in all-in-view, four to 12 (depending on the number of visible satellites) independent timing solutions are available. It has been observed that for extremely short baselines, all-in-view and common view should yield identical values. It was reported in [30] that for a baseline of 600 km between OP (France) and PTB (Germany), the uncertainty (σ) was 0.64 ns for all-in-view, and 0.66 ns for common view, showing hardly any difference between the two uncertainties. For longer baselines, the number of satellites that are jointly visible at both sites for common view decreases, and this decrease in the number of satellites degrades the common-view precision relative to all-in-view. For example [30], for a 17000 km baseline (i.e., between TL of Taipei and USNO of USA), the σ was 0.56 ns for all-in-view, but 1.42 ns for common view, exhibiting significant deterioration of uncertainty in common view. Furthermore, in all-in-view, a single pivot laboratory (e.g., PTB), as shown in Figure 12, may used to link all participating laboratories. In view of the above observations, the BIPM software has been modified to simultaneously perform the computation of TAI in both the common view and all-in-view modes.

9. Geodetic Receiver with Carrier-Phase Measurements in Common-View Mode

Geodetic GPS receivers have the advantage of additionally providing P-code and carrier-phase data on both frequencies L1 and L2, with a noise level significantly smaller than the noise on the C/A code. The use of a geodetic receiver for time and frequency transfer is based on a combined analysis of code and carrier-phase measurements, and using a consistent modeling of these

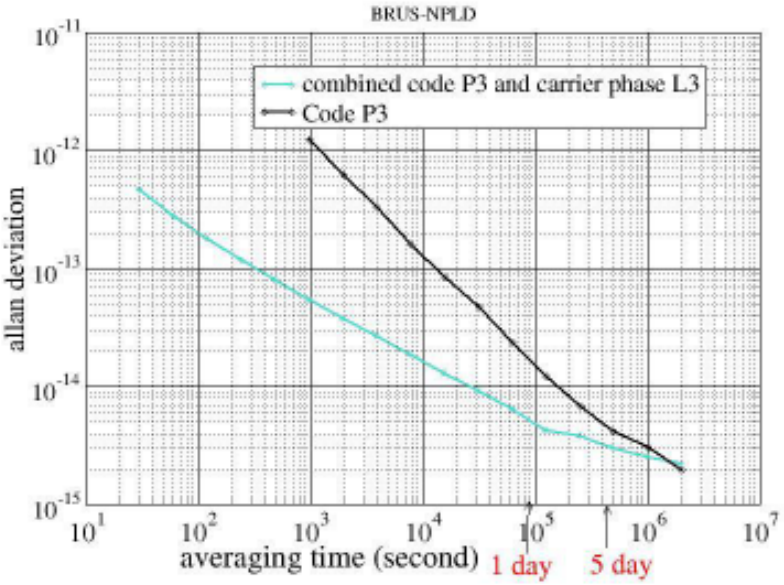

Figure 14. The improvement in the Allan deviation of time transfer by carrier phase over that by precise code [34]).

Table 5. The capabilities of GPS timing by various techniques.

Technique		Timing Uncertainty 1σ
Method I (Direct Access)		< 30 ns
Method II (Common View)	Single-Channel	<10 ns
	Multi-Channel	< 5 ns
	Carrier-Phase	< 500 ps

measurements. The noise level of the carrier-phase measurements is about 100 times smaller than the corresponding noise level on the code measurements. For this reason, the carrier phases have been used since 1980 for different geodetic applications requiring very high precision. During the last few years, the potential of GPS carrier phases for time transfer was recognized and demonstrated by different authors [34-42].

The carrier-phase observation ??(for L1 or L2) by a receiver, *A*, and a satellite, *i*, can be expressed by

$$\frac{\lambda}{2\pi}\varphi_A = \rho_A - c\left(\Delta t_A - \Delta t_s\right)$$

$$-\delta_{ion}(A)+\delta_{trop}(A)+\lambda N_A + noise\,, \quad (2)$$

where λ is the carrier wavelength ($\lambda/c = 635$ ps for L1), ρ is the range of the corresponding satellite, Δt_A and Δt_s are the receiver and satellite clock offsets, δ_{ion} and δ_{trop} are the ionospheric and tropospheric delays, and N_A is the cycle ambiguity, constant during one track if there is no cycle slip.

The clock offset between two receivers, *A* and *B*, (Figure 7) is obtained from the single difference

$$\frac{\lambda}{2\pi}(\varphi_A - \varphi_B) = \rho_A - \rho_B + c\left(\Delta t_{AB}\right)$$

$$+\left[\delta_{ion}(A)-\delta_{ion}(B)\right]-\left[\delta_{trop}(A)-\delta_{trop}(B)\right]$$

$$+\lambda\left(N_A - N_B\right), \quad (3)$$

where Δt_{AB} is the clock offset between the two receivers *A* and *B*.

A comparison of the results obtained from code observations and carrier-phase analysis is shown in Figures 13 and 14. It is of interest to note that improvements in the Allan deviation by the use of carrier phases for lower averaging times disappear for averaging times longer than five days. Due to the cycle ambiguities, it is not possible to determine the absolute value of Δt_{AB}. The absolute value of the clock offset can only be determined by combining the code information and the phase observations. Thus, the precision of the clock offset is limited by the code noise level of the receiver type, but may be smaller than 1 ns.

However, the relative frequency offset (*y* in the Appendix) being the rate of change of phase, the carrier-phase observable can thus be used for the determination of frequency without knowledge of cycle ambiguities. A very precise determination of the frequency transfer is thus possible, as it is affected only by the small noise level of the carrier phase.

10. Concluding Remarks

GPS reached the program's requirements for full operational capability (FOC) by 1995, and completed its original design goals. However, additional advances in technology and new demands on the existing system led to the effort to modernize the GPS system. The modernization program [60] aims to improve the accuracy and availability for all users. It involves new ground stations, new satellites, and four additional navigation signals. The new civilian signals are called L2C, L5, and L1C; the new military code is called the M-code. The entire program is expected to be fully operational by 2011.

The current status of GPS timing capabilities is summarized in Table 5. Continuous efforts have been going on to counter the identified sources of error. Among the hardware considerations, one has to account for hardware delays, which have to be determined by calibration. These hardware delays are sensitive to the ambient temperature variations. This is the reason why it is important to have temperature stabilization in the laboratories. Some experiments were performed to test the sensitivity of the equipment (receivers, amplifiers, etc.) to temperature variations, and the results showed the importance of keeping the temperature constant within 1° C. Concerning the antenna, the experiments showed maximum diurnal variations of 40 ps for the carrier phases, while up to 2 ns for the code measurements. Some antenna cables exist with a very small sensitivity to temperature variations, and these should be used for precise time and frequency transfer.

The TAI has very good long-term stability, but is available only after several weeks. With a one-day delay, the IGS provides a time scale, IGST (or IGRT, for the rapid version), with very good short-term stability. The IGS time scale is realized from the IGS clock combinations – i.e., satellite and receiver clock offsets based on time transfer between IGS receiver and/or satellite clocks – and it is computed from the combination of code and phase observations. On a daily basis, the IGS provides the clock synchronization errors of all station clocks with respect to IGST. This tool can be used by IGS stations for short-term monitoring of their clocks. The IGS time scale is steered to TAI in order to ensure long-term stability.

The GPS common-view method (and also, the all-in-view method) have become the main tools for international time and frequency transfer, because of the good cost performance of user equipment. However, due to the development of precise and accurate frequency standards, such as Cs fountain frequency standards [61], more-accurate and precise time- and frequency-transfer techniques have become necessary. Based on such demands, several time and frequency institutes have started to research and develop the two-way satellite time- and frequency-transfer (TWSTFT) system [44-50]. A working group on TWSTFT, under the CCTF, was established to discuss the technical issues and operation of the TWSTFT link. Nowadays, TWSTFT has become another tool for international time and frequency transfer for TAI calculation.

GNSS (Global Navigation Satellite System) is the common name for a global satellite-based system for navigation, positioning, and time transfer. In addition to GPS, Russia also has a similar system, called GLONASS (GLObal Navigation Satellite System) [51-56]. A European effort is underway to field a GNSS system named Galileo. GNSS also includes a number of GPS-enhancing systems that are being prepared for implementation. In the United States, the wide area augmentation system (WAAS) has two operational broadcasting satellites. The European Geostationary Navigation Overlay Service (EGNOS) is in the testing phase, and is actually used operationally by the European community [58]. The Japanese Quasi-Zenith Satellite System (QZSS) is steadily approaching the launch stage [59]. As more GNSS systems and their augmentations are fielded, the use of GNSS satellites is expected to bring more reliability and some further improvements in the performance of time links. The harmonization of reference time scales used by different GNSS systems is essential for optimum utilization of total GNSS. This can only be achieved by technical efforts coupled with political cooperation.

11. Appendix

The instantaneous output of an oscillator in the presence of noise may be represented by

$$V(t) = V_0 + \xi(t)\sin\left[2\pi f_0 t + \theta(t)\right]. \quad (4)$$

The instantaneous frequency of the output voltage is thus equal to

$$f(t) = f_0 + \frac{1}{2\pi}\frac{d\theta}{dt} = f_0 + \Delta f(t). \quad (5)$$

The normalized (dimensionless) instantaneous frequency offset from the nominal value, f_0, is

$$y(t) = \frac{\Delta f}{f_0} = \frac{1}{2\pi f_0}\frac{d\theta}{dt}. \quad (6)$$

Another useful quantity is the time integral of $y(t)$:

$$x(t) = \int_0^t y(t)dt\,\frac{\Delta f}{f_0} = \frac{\theta(t)}{2\pi f_0}. \quad (7)$$

When one measures the phase difference between the 1 pps outputs of two timing systems (in dimensions of time) with the help of a time-interval counter (TIC), one actually records $x(t)$. This is related to the normalized relative instantaneous frequency offset, $y(t)$, by

$$y(t) = \frac{\partial x}{\partial t}. \quad (8)$$

$x(t)$ is proportional to the instantaneous phase, but has the dimension of time. $x(t)$ is also sometimes termed phase-time. Existing methods of measurement do not allow the measurement of an instantaneous sample of $y(t)$. The result of a frequency measurement is always obtained from an average over a finite time interval, τ:

$$\bar{y}_k(t_k,\tau) = \frac{1}{\tau}\int_{t_k}^{t_{k+\tau}} y(t)dt = \frac{1}{\tau}\left[x(t_k+\tau) - x(t_k)\right] \quad (9)$$

The terms $x(t_k+\tau)$ and $x(t_k)$ are proportional to the instantaneous phase (time) difference obtained from the comparison between two clocks at date t_k and $t_{k1} = t_k + \tau_0$, respectively.

The time-domain characterization [57] of the frequency stability is through the Allan variance or two-sample zero-dead-time variance (i.e., Allan variance (σ_y^2)) , which is the most common way to quantify the random frequency stability of an oscillator in the time domain. σ_y *is also called the Allan deviation*. σ_y^2 is normally defined as

$$\sigma_y^2(\tau) = \frac{1}{2}\left\langle(\bar{y}_{k+1} - \bar{y}_k)^2\right\rangle, \quad (10)$$

where $\langle\cdots\rangle$ denotes the average over a large number of samples, and $k = 1, 2, 3, \ldots$.

The second differences are used to calculate the Allan variance. It may be easily shown from Equation (9) that

$$\bar{y}_{k+1} - \bar{y}_k = \frac{1}{\tau}\left[x(t_k+2\tau) - 2x(t_k+\tau) + x(t_k)\right]. \quad (11)$$

So, by combining Equations (10) and (11), one may find the Allan deviation directly from the measurement of $x(t)$.

12. References

1. C. Audoin and J. Vervier, *Physics of Atomic Frequency Standards*, London, Adam Higler, 1988.
2. P. Kartaschoff, *Frequency and Time*, New York, Academic Press, 1978.
3. J. L. Jespersen, B. E. Blair, and L. E. Gatterer, "Characterization and Concept of Time and Frequency Dissemination," *Proc. IEEE*, **30**, 5, May 1972, pp. 502-521.
4. P. Banerjee, M. Saxena, and B. S. Mathur, "Microprocessor Based Generation of Time and Frequency Signal Format for HF Broadcast," *Measurement Science and Technology*, **1**, 5, 1990, pp. 934-938.
5. Cyrus E. Potts and Bernard Wieder, "Precise Time and Frequency Dissemination via Loran-C System," *Proc. IEEE*, **60**, 5, May 1972, pp. 530-539.
6. J. Tolman, V. Ptacek, A. Soucekand and R. Stecher, "Microsecond Clock Comparison by Means of TV Synchronizing Pulses," *IEEE Trans. Instr. Meas*., **IM-16**, September 1967, pp. 247-254.
7. B. Z Kovacevic, B. R. Dimitrijevic, M. Z. Arsic and N. S. Kovacevic, "Precise Real Time Signal Dissemination Over the TV Broadcasting Satellite," *Radio Sci*., **14**, 4.
8. P. Banerjee, M. Saxena, and B. S. Mathur, "Passive TV Technique for Time Transfer via Indian Satellite INSAT-1B," *IEEE Trans. Inst. Meas.*, **IM-36**, 2, June 1987, pp. 579-583.
9. R. E. Bechler, D. D. Davis, and J. B. Milton, "GOES Satellite Time Code Dissemination: Description and Operation," NBS Publications, January 1988, pp. 250-300.
10. A. Sen Gupta, A. K. Hanjura, and B. S. Mathur, "Satellite Broadcasting of Time and Frequency Signals," *Proc. IEEE*, **79**, 7, July 1991, pp. 973-982.
11. B. Hoffmann-Wellenhof, H. Lichtenneger and J. Collins, *GPS – Theory and Practice*, New York, Springer Wein, 1997.
12. Michael A. Lombardi, "Traceability in Time and Frequency Metrology," *Cal Lab: The International Journal of Metrology*, September-October 1999, pp. 33-40.
13. W. Lawandowski and C. Thomas, "GPS Time Transfer," *IEEE Proc.*, **79**, 1991, pp. 991.
14. W. Lewandowski, J. Azoubib, W. J. Klepczynski, "GPS: Primary Tool for Time Transfer," *Proceedings of the IEEE*, **87**, 1, 1999, pp. 163-172.
15. Richard Hambly, Thomas A. Clark, "Critical Evaluation of the Motorola M12+GPS Timing Receiver vs. the Master Clock," *Proc. 34th Annual PTTI Meeting*, 2004, pp. 109-115.
16. F. Cordara and V. Pettiti, "Short Term Characterization of GPS Disciplined Oscillators and Field Trial for Frequency of Italian Calibration Centres," *Proc. IEEE Frequency Control Symposium*, April 1999, pp. 404-407.
17. J. A. Davis and J. M. Furlong, "Report on the Study to Determine the Suitability of GPS Disciplined Oscillators as Time and Frequency Standards Traceable to the UK National Time Scale UTC (NPL)," Nat. Phys. Lab. (NPL) Report CTM 1, 1997 .
18. P. Banerjee, Suman, A. K. Suri, Arundhati Chatterjee and Anindya Bose, "A Study on the Potentiality of the GPS Timing Receiver for Real Time Applications," *Mease. Sci. Technol.*, **18**, 2007, pp. 3811-3815.
19. R. B. Langley, "The GPS Error Budget," *GPS World*, **8**, 3, 1997, pp. 51-56.
20. J. A. Klobuchar, "Ionospheric Time-delay Algorithm for Single Frequency GPS Users," *IEEE Transactions on Aerospace and Electronic Systems*, **AES-23**, 3, 1987, pp. 325-331.
21. J. Aarons, "The Longitudinal Morphology of Equatorial F-layer Irregularities Relevant to Their Occurrence," *Space. Sci. Rev.*, **63**, 1993, pp. 209.
22. P. Banerjee, Anindya Bose and Ashish Dasgupta, "Effect of Scintillation on Timing Application of GPS in Indian Subcontinent," *IEEE Trans. Instrm. Meas.*, **56**, 5, October 2007, pp. 1596-1600
23. S. Basu and Su. Basu, "Equatorial Scintillations – A Review," *J. Atoms. Terr. Phy*., **43**, 1985, pp. 473.
24. T. Bandyopadhyay, A. Guha, S. DasGupta, P. Banerjee and A. Bose, "Degradation of Navigational Accuracy with Global Positioning System During Periods of Scintillation at Equatorial Latitudes," *Electronics Letters*, **33**, 1997, pp. 1010.
25. P. Banerjee, G. K. Goel, and B. S. Mathur, "Monitoring of GPS Signals at NPL, New Delhi for Precise Time Comparison," *Indian J. Radio & Space Phys*, **23**, 1994, pp. 416-420.
26. D. W. Allan and M. A. Weiss, "Accurate Time and Frequency Transfer During Common-View of a GPS Satellite," *Proc. Freq. Cont. Symp.*, 1980, pp. 334-346.
27. D. Matsakis, F. Arias, A. Bauch, J. Davis, T. Gotoh, M. Hosokawa, and D. Piester, "On Optimizing the Configuration of Time-Transfer Links Used to Generate Tai," *Proc. of the 20th European Time and Frequency Forum*, 2006.
28. D. W. Allan, D. D. Davis, M. A. Weiss, A. J. Clements, B. Guinot, M. Granveaud, K. Dorenwendt, B. Fischer, P. Hetzel, S. Aoki, M.-K. Fujimoto, L. Charron, and N. Ashby, "Accuracy of International Time and Frequency Comparisons Via Global Positioning System Satellites in Common-View," *IEEE Trans. Instr. Meas.*, **34**, 1985, pp. 118-125.
29. W. Lewandowski, D. Matsakis, G. Panfilo and P. Tavella, "The Evaluation of Uncertainties in [UTC – UTC (k)]," *Metrologia*, **43**, 2006, pp. 278-276.
30. G. Petit and Z. Jiang, "GPS All in View Time Transfer for TAI Computation," *Metrologia*, **45**, February, pp. 35-45.
31. K. Senior, P. Koppang and J. Ray, "Developing on IGS Time Scale," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, **50**, 6, June 2003, pp. 585-593.
32. P. Banerjee and D. Matsakis, "Time Transfer through GPS, and the Harmonization of GPS, GLONASS and Galileo for Timing," *MAPAN: J. Metrology Society of India*, **21**, 4, December 2006, pp. 229-242.
33. Judah Levine, "Time Transfer Using Multi-Channel GPS Receivers," *IEEE Transactions of Ultrasonics, Ferroelectrics, and Frequency Control*, **46**, 2, 1999, pp. 392-398.
34. P. Defraigne, P. Banerjee, and W. Lewandowski, "Time Transfer through GPS," *Indian J. Radio & Space Phys.*, **36**, August 2007, pp. 303-312.
35. G. Petit and Z. Jiang, "Stability and Accuracy of GPS P3 Time Links," *Proceedings of the 34th Precise Time and Time Interval (PTTI) Meeting*, 2004, pp. 31-39.
36. R. Dach, T. Schildknecht, U. Hugentobler, L.G. Bernier, and G. Dudle, "Continuous Geodetic Time Transfer Methods," *IEEE Trans. Ultrason., Ferroelect., and Freq. Contr.*, **53**, 7, July 2006, pp. 1250- 1259.
37. T. Schildknecht, G. Beutler, W. Gurtner, and M. Rothacher, "Towards Subnanosecond GPS Time Transfer using Geodetic Processing Techniques," *Proc. of the 4th Eur. Freq. Time Forum, EFTF 90*, 1990, pp. 335-346.
38. K. Larson, J. Levine, L. Nelson, and T. Parker, "Assessment of GPS Carrier-Phase Stability for Time-Transfer Applications," *IEEE Trans. Ultrason., Ferroelect., and Freq. Contr.*, **47**, 2, 2000, pp. 484-494.
39. D. Matsakis, K. Senior, and P. Cook, "Comparison of Continuously Filtered GPS Carrier Phase Solutions with Independent Daily GPS Carrier Phase Solutions and with Two-Way Satellite Time Transfer," *Proc. of the 33rd Precise Time and Time Interval (PTTI) Meeting*, 2001, pp. 63-88.
40. P. Defraigne and G. Petit, "Time Transfer to TAI Using Geodetic Receivers," *Metrologia*, **40**, 2003, pp. 184-188.
41. C. Bruyninx, P. Defraigne, "On the Link Between GPS Pseudorange Noise and Day-Boundary Discontinuities in Geodetic Time Transfer Solutions," *GPS solutions*, 2007, in press.
42. Z. Jiang, G. Petit, P. Defraigne, "Combination of GPS Carrier Phase Data with a Calibrated Time Transfer Link," EFTF 2007.

43. Kristine Larson and Judah Levine, "Time Transfer Using the Phase of the GPS Carrier," *IEEE Trans. on Ultrason. Ferroelect. and Freq. Contr.*, **45**, 3, May 1998, pp. 539-540.
44. M. Brunet, "Synchronization of Atomic Clocks through the 'Symphonie' Satellite," *Radio Science*, **14**, 4, 1979, pp.721-730.
45. B. S. Mathur, P. Banerjee, P. C. Sood, M. Saxena, N. Kumar, and A. K. Suri, "Precise Time and Frequency Intercomparison between NPL, India and PTB, Federal Republic of Germany via Satellite Symphonie-1," *Proc. 12th Annual PTTI Meeting*, November 1980, pp. 863-875.
46. Y. Saburi, M. Yamamoto, and K. Harada, "High-Precision Time Comparison via Satellite and Observed Discrepancy of Synchronization," *IEEE Trans. on Instr. and Meas.*, **IM-25**, 4, December 1976, pp. 473-477.
47. M. Imae, H. Okazawa, T. Sato, M. Uratsuka, et al., "Time Comparison Experiments with Small K-band Antennas and SSRA Equipments via a Domestic Geostationary Satellite," *IEEE Trans. Instr. and Meas.*, **IM-32**, 1, March 1983, pp. 199-203.
48. D. Kirchner, H. Ressler, P. Grudler, F. Baumont, Ch. Veillet, W. Lewandowski, W. Hanson, W. Klepczynski and P. Uhrich, "Comparison of GPS Common-View and Two Way Satellite Time Transfer Over a Baseline of 800 km," *Metrologia*, **30**, 1993, pp. 813.
49. P. Hartl, N. Gieschen, K. M. Mussener, W. Schafer, and C. M. Wende, "High Accuracy Global Time Transfer via Geosynchronous Telecommunication Satellite with MITREX," *J. of Flight Scie. and Spac. Rese.*, **7**, 1983, pp. 335-342.
50. A. Bauch, J. Achkar, S. Bize, D. Calonico, R. Dach, R. Hlava, L. Lorini, T. Parker, G. Petit, D. Piester, K. Szymaniec and P. Uhrich, "Comparison Between Frequency Standards in Europe and the USA at the 10^{-15} Uncertainty Level," *Metrologia*, **43**, 2006, pp. 109-120.
51. P. Banerjee, Anindya Bose, and Ashish Dasgupta, "The Usefulness of GLONASS for Positioning in the Presence of GPS in Indian Subcontinent," *J. of Navigation (UK)*, **55**, 3, 2002, pp. 463-475.
42. J. Gouzhva et al., "High-Precision Time and Frequency Dissemination with GLONASS," *GPS World*, July/August 1992, pp. 40-49.
53. W. Lewandowski and J. Azoubib, "GPS+GLONASS: Toward Sub-Nanosecond Time Transfer," *GPS World*, November 1998, pp. 30-39.
54. J. Azoubib, and W. Lewandowski, "Test of GLONASS Precise-Code Time Transfer," *Metrologia*, **37**, 2000, pp. 55-59.
55. W. Lewandowski, J. Nawrocki, and J. Azoubib, "First Use of IGEX Precise Ephemerides for International GLONASS P-Code Time Transfer," *Journal of Geodesy*, **75**, 2000, pp. 620-625.
56. W. Lewandowski, J. Azoubib, G. de Jong, J. Nawrocki and J. Danaher "A New Approach to International Time and Frequency Comparisons: "All-In-View" Multi-Channel GPS+GLONASS Observations," *Proc. ION GPS-97*, 1997, pp. 1085-1091.
57. D. W. Allan, "Time and Frequency (Time Domain) Characterization, Estimation and Prediction of Precision Clocks and Oscillators," *IEEE Trans. on Ultrason. Ferroelect. and Freq. Contr.*, 34, 1984, pp. 647.
58. http://en.wikipedia.org/wiki/European_Geostationary_Navigation_Overlay_System.
59. M. Hosokawa, Yukio Takahashi, Shin'ichi Hama, Hiroshi Toriyama, and Takao Morikawa, "Overview of Research Activities on Time and Frequency at the National Institute of Information and Communications Technology," *Proceedings of the 36th Precise Time and Time Interval (PTTI) Meeting*, 2004, pp. 196-206.
60. Richard D. Fontana, Wai Cheung, and Tom Stansell, "The Modernized L2 Civil Signal, Leaping Forward in the 21st Century," *GPS World*, September 2001.
61. Steven R. Jefferts, Thomas P. Heavner and Elizabeth A. Donley, "Cesium Primary Frequency References," *MAPAN : J. Metrology Society of India*, **21**, 4, December 2006, pp. 191-199.
62. E. F. Arias, "International Time Scales BIPM Activities and Recent Progress," talk delivered at NPLI, New Delhi, India, August 31, 2007.

Conferences

CONFERENCE REPORT

12TH INTERNATIONAL SYMPOSIUM ON EQUATORIAL AERONOMY (ISEA 2008)

Crete, Greece, 18 - 24 May 2008

The International Symposium on Equatorial Aeronomy (ISEA) is a historic meeting, initiated by ionospheric scientists in the early sixties soon after the Jicamarca radar started its operation near Lima, Peru. ISEA has been established over the years as an important event for the world's research community interested in the physics of the low- and mid-latitude upper atmosphere and ionosphere. Since the first symposium in Huaychulo, Peru nearly five decades ago, ISEA is held regularly every 3 to 5 years in different locations around the globe. It represents an opportunity for researchers in the aeronomic community to review and evaluate their scientific achievements over the period since the previous ISEA, share their most recent results and ideas, and discuss possibilities for new directions in research, joint experiments and observational campaigns.

The 12th International Symposium on Equatorial Aeronomy (ISEA-12), was hosted by the Ionospheric Physics Lab, at the Department of Physics, University of Crete. It was held from 18 to 24, May 2008 at the Royal Knossos Conference Hotel to the east of Heraklion on the island of Crete, Greece. 170 participants, both senior and young scientists, from 25 countries, attended the symposium. A total of 250 papers (about 150 oral and 100 posters) were presented by invited and contributing authors during 11 scientific sessions distributed over a period of a week. The topics covered a wide range of research areas, reflecting the need to study the Earth's ionosphere - atmosphere system in a coupled sense. ISEA-12 comprised sessions on the dynamics of the middle atmosphere, mesosphere and thermosphere, *E*- and *F*-region plasma physics and ionospheric electrodynamics, including large scale ionospheric modeling and simulation, atmosphere-ionosphere coupling processes and phenomena, magnetic storm and space weather effects, and a session on new experimental techniques and instruments. In addition, and for the first time in its long history, ISEA-12 started with a full day of tutorials on key topics given by leading members of the aeronomic community, and ended with a session of invited expert talks on future research trends and unresolved problems. The works of the symposium, its program and abstracts, are included in the ISEA-12 Book of Abstracts which is available for inspection and download at the symposium's web page: http://isea12.physics.uoc.gr .

The ISEA-12 Organizing Committee is also proud to have produced the ISEA-12 Book of Tutorials, a high quality publication volume of 150 pages that was offered to all the symposium participants. Its scope was to allow the reader to gain a wider perspective on fundamental scientific aspects of low- and middle-latitude aeronomy, as they have evolved over the past five decades of research since the first ISEA meeting. It includes six tutorials which constitute comprehensive reviews of the present state of knowledge on key research areas of ionospheric and upper atmospheric science. The topics covered in this book are: upper atmosphere waves and dynamics (Prof. Bob Vincent, University of Adelaide), equatorial E- and F-region plasma irregularities and instabilities (Prof. Don Farley, Cornell University; Dr. Ron Woodman, Radio Observatorio Jicamarca), midlatitude electrodynamics and plasma physics (Prof. Michael Kelley, Cornell University), internal and external influences on ionospheric electrodynamics at low and middle latitudes (Prof. Rod Heelis, University of Texas at Dallas), as well as lower and middle atmosphere electrical phenomena and electrodynamics (Prof. Umran Inan, Stanford University). The emphasis is on the observational characteristics of the various aeronomic phenomena at low and middle latitudes and the governing physical principles, theories and mechanisms, which define a working set of interpretations. The material of this book aims to allow the members of the community, whether they are specialists or not, to comprehend the recent developments in research and thus obtain a clearer picture of the current state of understanding concerning low- and middle-latitude aeronomy. The organizing committee and the participants of ISEA-12 dedicated this book to the memory of Tor Hagfors, an outstanding scientist, a dear colleague and friend, who suddenly passed away on January 17, 2007. Based on the kind approval of the tutorial authors, the Organizing Committee makes the ISEA-12 Book of Tutorials available to everyone interested, by posting its entire pdf version in the symposium's web page at http://isea12.physics.uoc.gr .

The works (including also most of the oral and poster presentations in ppt or pdf forms) and other happenings of ISEA-12 will be compiled in a user-friendly DVD volume, presently under creation, to be distributed to every

participant. Also a large number of authors have expressed interest in publishing their presentations in a Special Issue of *Annales Geophysicae* entitled "12th International Symposium on Equatorial Aeronomy (ISEA-12)". The team of Guest Editors includes : Christos Haldoupis (University of Crete, Greece), Jonathan Makela and Erhan Kudeki (University of Illinois at Urbana-Champaign, USA), Jorhe Chau (Radio Observatorio de Jicamarca, Peru), Dora Pancheva (Academy of Sciences Geophysical Institute, Bulgaria), and Dave Hysell (Cornell University, USA). The paper submission deadline is September 30, 2008.

The ISEA-12 organizing committees, who have been responsible for the scientific program and the overall organization, wish to cordially thank everyone for their attendance and valuable support of the symposium in every respect and particularly for their scientific contributions. Many thanks are also extended to the national (Greek) and international sponsors (http://isea12.physics.uoc.gr), including the International Union of Radio Science (URSI), for providing funds that assisted the participation of many colleagues from various countries, mostly young scientists and PhD students.

At the end of ISEA-12 it was unanimously felt that the symposium had been a great success, both scientifically and socially, and that the low- and mid-latitude aeronomic community has gained new momentum to carry on its research tasks and challenges into the future. To maintain the momentum and progress, the ISEA-12 Science Advisory Committee accepted the invitation of our Peruvian colleagues to host the 13th International Symposium on Equatorial Aeronomy in Peru in 2012. This will commemorate the 50th anniversary of ISEA. The journey continues!

Christos Haldoupis, Chair, ISEA-12
Email:chald@physics.uoc.gr

CONFERENCE ANNOUNCEMENT

VIIITH INTERNATIONAL SYMPOSIUM AND EXHIBITION ON ELECTROMAGNETIC COMPATIBILITY AND ELECTROMAGNETIC ECOLOGY (EMC 2009)

Sint-Petersburg, Russia, 16 - 19 June 2009

Objective

The 8th International Symposium on Electromagnetic Compatibility and Electromagnetic Ecology (EMC&EME) will offer a new opportunity for scientists, engineers and students working in the area of EMC to present the progress in their work and to discuss problems of current mutual interest.

The Symposium traditionally takes place at St. Petersburg State Electrotechnical University "LETI", St. Petersburg, Russia. The Symposium EMC'2009 will be held in June in the fine period of White Nights. This is a pleasant time to get acquainted with the architecture and museums of the city (the Hermitage, the Russian museum and others), for walks around the city and for visiting famous palaces (Peterhof, Pavlovsk, Pushkin).

Topics

1. Theoretical problems of EMC&EME
2. EMC of radio-electronic equipment
3. Spectrum management and monitoring
4. EMC in electrical engineering and power systems
5. EMC and EME specifically for mobile objects (vessels, airplanes, railway transport and others)
6. Research of natural electromagnetic radiations
7. Equipment design with regard to EMC&EME, technology, materials and components
8. Electromagnetic monitoring, measurement, certification and test equipment
9. EME problems: influence of electromagnetic radiation on biological objects, allowable norms of radiation and ecological protection
10. EMC&EME education

Contact

Discone-Centre Ltd. at St. Petersburg State
Electrotechnical University "LETI":
Tel: +7 812 234-48-40
Fax: +7 812 234-46-81
E-mail: discone@mail.wplus.net
http://www.eltech.ru/emc

URSI CONFERENCE CALENDAR

URSI cannot be held responsible for any errors contained in this list of meetings.

July 2008

COSPAR 2008 - 37th Scientific Assembly of the Committee on Space Research and Associated Events "50th Anniversary Assembly"
Montreal, Canada, 13 - 20 July 2008
cf. Announcement in the Radio Science Bulletin of March 2007, P. 58
Contact : COSPAR Secretariat, c/o CNES, 2 place Maurice Quentin, 75039 Paris Cedex 01, France, Tel: +33 1 44 76 75 10, Fax: +33 1 44 76 74 37, E-mail : cospar@cosparhq.cnes.fr, Web : http://www.cospar2008.org

EUROEM 2008 - European Electromagnetics
Lausanne, Switzerland, 21-25 July 2008
Contact : EUROEM'08, EPFL-STI- LRE, Station 11, CH-1015 Lausanne, Switzerland, Tel : +41-21-693 26 20, Fax : +41-21-693 46 62, E-mail: information@euroem.org, Web : http://www.euroem.org

August 2008

URSI GA08 - XXIXth URSI General Assembly
Chicago, IL, USA, 9-16 August 2008
Contact : URSI Secretariat, c/o INTEC, Ghent University, Sint-Pietersnieuwstraat 41, B-9000 Ghent, Belgium, Tel. : +32 9 264 3320, Fax : +32 9 264 4288, E-mail : info@ursi.org Web: http://www.ece.uic.edu/2008ursiga

ICWCUCA - Second International Conference on Wireless Communications in Underground and Confined Areas
Val-d'Or, Canada, 25-27 August 2008
Contact : Laboratoire de recherche Télébec Mobilité en communications souterraines (LRTCS), International Conference on Wireless Communications in Underground and Confined Areas (ICWCUCA), 450, 3e avenue, local 103, Val-d'Or (Québec) J9P 1S2, Canada, Tel. +1 819 874-7400, Fax: +1 819 874-7166, E-mail: lrcs@uqat.ca , Web: http://www.icwcuca.ca/

September 2008

EMC Europe 2008 - International Symposium on Electromagnetic Compatibility 2008 and Technical Exhibition
Hamburg, Germany, 8-12 September 2008
Contact: EMC Europe 2008 Secretariat, c/o Dr. J. L. ter Haseborg, Technische Universitaet Hamburg-Harburg, Institut fuer Messtechnik - EMV, Harburger Schlossstrasse 20, D-21079 Hamburg, Germany, Fax: +49 40-42878-2382, E-mail: info@emceurope2008.org , Web: http://www.emceurope2008.org

Metamaterials 2008 - 2nd International Congress on Advanced Electromagnetic Materials in Microwaves and Optics
Pamplona, Spain , 21-26 September 2008
Contact: http://congress2008.metamorphose-vi.org/

November 2008

Microwave 2008 - International Conference on Recent Advances in Microwave Theory and Applications
Jaipur, India, 21 - 24 November 2008
Contact: Dr. Deepak Bhatnagar, Department of Physics, University of Rajasthan, Jaipur, 302004, India, Fax +91 1412 702645, E-mail: dbhatnagar_2000@rediffmail.com , Web: www.uniraj.ernet.in/~microwave

May 2009

URBAN 2009 - Data fusion and Remote Sensing in Urban areas
Shanghai, China, May 2009
Contact : Shanghai Association for Science and Technology (SAST), No.47 Nanchang Road, Shanghai 200020, China (SAST), Tel: 86-21-6358 0841-207, Fax: 86-21-6327 1590, E-mail: urban-urs2009@163.com, Web: http://www.urban-remote-sensing-2009.org.cn

June 2009

EMC-2009 - VIII International Symposium and Exhibition on Electromagnetic Compatibility and Electromagnetic Ecology
St. Petersburg, Russia, 16-19 June 2009
Contact : St.Petersburg State Electrotechnical University "LETI" , 5, Prof. Popov Street, St. Petersburg, 197376, Russia, Underground Station "PETROGRADSKAYA", Phone: +7 812 234-48-40, Fax: +7 812 234-46-81, E-mail: discone@mail.wplus.net, Web: http://www.eltech.ru/emc

URSI cannot be held responsible for any errors contained in this list of meetings.

News from the URSI Community

NEWS FROM A MEMBER COMMITTEE

FRANCE
Médaille du CNFRS

La Médaille du Comité National Français de Radioélectricité Scientifique (CNFRS) section française de l'Union Radio Scientifique Internationale (URSI) a été décernée le 20 mai 2008 à Pierre Baüer. Le CNFRS souhaite ainsi souligner l'importance des contributions de Pierre Baüer au développement des connaissances de l'ionosphère, à la compréhension des interactions entre couches denses de l'atmosphère et champs électromagnétiques, à la caractérisation des calottes glaciaires, comme animateur des communautés scientifiques de ces domaines et plus généralement comme responsable de l'URSI.

Cette médaille est « destinée à honorer des personnes qui ont œuvré pour le renom de la Science en Radioélectricité et/ou participé d'une manière très significative à la vie et au renom du CNFRS/URSI.

Pierre Baüer, ingénieur électrotechnicien de l'Institut national polytechnique de Grenoble, entreprend en 1964 une activité de recherche en physique spatiale dans un laboratoire de la NASA à l'université du Michigan. Son travail de thèse (*PhD*) porte sur l'étude du plasma de l'environnement terrestre. Sur proposition de Michel Petit, il rejoint l'équipe du sondeur ionosphérique à diffusion incohérente du CNET en 1969, en qualité de chercheur CNRS.

Après un nouveau séjour aux États-Unis (1974-1975) en tant que *Senior post doctoral fellow* de la *National Academy of Sciences*, il assume successivement les responsabilités de directeur adjoint du CRPE/CNET, de directeur du Service d'aéronomie, de directeur du Centre d'études spatiales de la biosphère et de directeur adjoint de la recherche de Météo France.

Ses recherches concernent, dans une première phase, l'étude expérimentale de l'ionosphère, des phénomènes auroraux et de la thermosphère grâce à la technique de la diffusion incohérente des ondes électromagnétiques par l'ionosphère. Elles évoluent ensuite vers la télédétection des couches denses de l'atmosphère et des calottes glaciaires.

Il est vice-président, puis président de la commission G (Radioélectricité ionosphérique et propagation) de l'URSI de 1978 à 1984, président du Conseil scientifique d'EISCAT (*European Incoherent Scatter*) de 1979 à 1982, membre du conseil d'administration d'EISCAT de 1982 à 1990, coordinateur du programme scientifique des Assemblées générales de l'URSI de 1987 et 1990, trésorier de l'URSI de 1990 à 1993, président du CNFRS de 1990 à 1992, président de l'URSI de 1993 à 1996, président sortant de l'URSI de 1996 à 1999, trésorier du CNFRS/URSI-France de 2001 à 2005.

The Journal of Atmospheric and Solar-Terrestrial Physics

SPECIAL OFFER TO URSI RADIOSCIENTISTS

AIMS AND SCOPE

The *Journal of Atmospheric and Terrestrial Physics* (JASTP) first appeared in print in 1951, at the very start of what is termed the "Space Age". The first papers grappled with such novel subjects as the Earth's ionosphere and photographic studies of the aurora. Since that early, seminal work, the Journal has continuously evolved and expanded its scope in concert with - and in support of - the exciting evolution of a dynamic, rapidly growing field of scientific endeavour: the Earth and Space Sciences. At its Golden Anniversary, the now re-named *Journal of Atmospheric and Solar-Terrestrial Physics* (JASTP) continues its development as the premier international journal dedicated to the physics of the Earth's atmospheric and space environment, especially the highly varied and highly variable physical phenomena that occur in this natural laboratory and the processes that couple them. The *Journal of Atmospheric and Solar-Terrestrial Physics* is an international journal concerned with the inter-disciplinary science of the Sun-Earth connection, defined very broadly. The journal referees and publishes original research papers, using rigorous standards of review, and focusing on the following: The results of experiments and their interpretations, and results of theoretical or modelling studies; Papers dealing with remote sensing carried out from the ground or space and with in situ studies made from rockets or from satellites orbiting the Earth; and, Plans for future research, often carried out within programs of international scope. The Journal also encourages papers involving: large scale collaborations, especially those with an international perspective; rapid communications; papers dealing with novel techniques or methodologies; commissioned review papers on topical subjects; and, special issues arising from chosen scientific symposia or workshops. The journal covers the physical processes operating in the troposphere, stratosphere, mesosphere, thermosphere, ionosphere, magnetosphere, the Sun, interplanetary medium, and heliosphere. Phenomena occurring in other "spheres", solar influences on climate, and supporting laboratory measurements are also considered. The journal deals especially with the coupling between the different regions. Solar flares, coronal mass ejections, and other energetic events on the Sun create interesting and important perturbations in the near-Earth space environment. The physics of this subject, now termed "space weather", is central to the Journal of Atmospheric and Solar-Terrestrial Physics and the journal welcomes papers that lead in the direction of a predictive understanding of the coupled system. Regarding the upper atmosphere, the subjects of aeronomy, geomagnetism and geoelectricity, auroral phenomena, radio wave propagation, and plasma instabilities, are examples within the broad field of solar-terrestrial physics which emphasise the energy exchange between the solar wind, the magnetospheric and ionosphericplasmas, and the neutral gas. In the lower atmosphere, topics covered range from mesoscale to global scale dynamics, to atmospheric electricity, lightning and its effects, and to anthropogenic changes. Helpful, novel schematic diagrams are encouraged. Short animations and ancillary data sets can also be accommodated. Prospective authors should review the *Instructions to Authors* at the back of each issue.

Complimentary Information about this journal:
http://www.elsevier.com/locate/JASTP?
http://earth.elsevier.com/geophysics

Audience:
Atmospheric physicists, geophysicists and astrophysicists.

Abstracted/indexed in:
CAM SCI Abstr
Curr Cont SCISEARCH Data
Curr Cont Sci Cit Ind
Curr Cont/Phys Chem & Sci
INSPEC Data
Meteoro & Geoastrophys Abstr
Res Alert

Editor-in-Chief:
T.L. Killeen, National Centre for Atmospheric Research, Boulder, Colorado, 80307 USA

Editorial Office:
P.O. Box 1930, 1000 BX Amsterdam, The Netherlands

Special Rate for URSI Radioscientists 2003:
Euro 149.00 (US$ 149.00)
Subscription Information
2002: Volume 65 (18 issues)
Subscription price: Euro 2659 (US$ 2975)
ISSN: 1364-6826

CONTENTS DIRECT:
The table of contents for this journal is now available pre-publication, via e-mail, as part of the free ContentsDirect service from Elsevier Science. Please send an e-mail message to cdhelp@elsevier.co.uk for further information about this service.

For ordering information please contact Elsevier Regional Sales Offices:

Asia & Australasia/ e-mail: asiainfo@elsevier.com
Europe, Middle East & Africa: e-mail: nlinfo-f@elsevier.com
Japan: Email: info@elsevier.co.jp
Latin America : e-mail: rsola.info@elsevier.com.br
United States & Canada : e-mail: usinfo-f@elsevier.com

Information for authors

Content

The *Radio Science Bulletin* is published four times per year by the Radio Science Press on behalf of URSI, the International Union of Radio Science. The content of the *Bulletin* falls into three categories: peer-reviewed scientific papers, correspondence items (short technical notes, letters to the editor, reports on meetings, and reviews), and general and administrative information issued by the URSI Secretariat. Scientific papers may be invited (such as papers in the *Reviews of Radio Science* series, from the Commissions of URSI) or contributed. Papers may include original contributions, but should preferably also be of a sufficiently tutorial or review nature to be of interest to a wide range of radio scientists. The *Radio Science Bulletin* is indexed and abstracted by INSPEC.

Scientific papers are subjected to peer review. The content should be original and should not duplicate information or material that has been previously published (if use is made of previously published material, this must be identified to the Editor at the time of submission). Submission of a manuscript constitutes an implicit statement by the author(s) that it has not been submitted, accepted for publication, published, or copyrighted elsewhere, unless stated differently by the author(s) at time of submission. Accepted material will not be returned unless requested by the author(s) at time of submission.

Submissions

Material submitted for publication in the scientific section of the *Bulletin* should be addressed to the Editor, whereas administrative material is handled directly with the Secretariat. Submission in electronic format according to the instructions below is preferred. There are typically no page charges for contributions following the guidelines. No free reprints are provided.

Style and Format

There are no set limits on the length of papers, but they typically range from three to 15 published pages including figures. The official languages of URSI are French and English: contributions in either language are acceptable. No specific style for the manuscript is required as the final layout of the material is done by the URSI Secretariat. Manuscripts should generally be prepared in one column for printing on one side of the paper, with as little use of automatic formatting features of word processors as possible. A complete style guide for the *Reviews of Radio Science* can be downloaded from http://www.ips.gov.au/IPSHosted/NCRS/reviews/. The style instructions in this can be followed for all other *Bulletin* contributions, as well. The name, affiliation, address, telephone and fax numbers, and e-mail address for all authors must be included with all submissions.

All papers accepted for publication are subject to editing to provide uniformity of style and clarity of language. The publication schedule does not usually permit providing galleys to the author.

Figure captions should be on a separate page in proper style; see the above guide or any issue for examples. All lettering on figures must be of sufficient size to be at least 9 pt in size after reduction to column width. Each illustration should be identified on the back or at the bottom of the sheet with the figure number and name of author(s). If possible, the figures should also be provided in electronic format. TIF is preferred, although other formats are possible as well: please contact the Editor. Electronic versions of figures *must* be of sufficient resolution to permit good quality in print. As a rough guideline, when sized to column width, line art should have a minimum resolution of 300 dpi; color photographs should have a minimum resolution of 150 dpi with a color depth of 24 bits. 72 dpi images intended for the Web are generally *not* acceptable. Contact the Editor for further information.

Electronic Submission

A version of Microsoft *Word* is the preferred format for submissions. Submissions in versions of T_EX can be accepted in some circumstances: please contact the Editor before submitting. *A paper copy of all electronic submissions must be mailed to the Editor, including originals of all figures.* Please do *not* include figures in the same file as the text of a contribution. Electronic files can be send to the Editor in three ways: (1) By sending a floppy diskette or CD-R; (2) By attachment to an e-mail message to the Editor (the maximum size for attachments *after* MIME encoding is about 7 MB); (3) By e-mailing the Editor instructions for downloading the material from an ftp site.

Review Process

The review process usually requires about three months. Authors may be asked to modify the manuscript if it is not accepted in its original form. The elapsed time between receipt of a manuscript and publication is usually less than twelve months.

Copyright

Submission of a contribution to the *Radio Science Bulletin* will be interpreted as assignment and release of copyright and any and all other rights to the Radio Science Press, acting as agent and trustee for URSI. Submission for publication implicitly indicates the author(s) agreement with such assignment, and certification that publication will not violate any other copyrights or other rights associated with the submitted material.

APPLICATION FOR AN URSI RADIOSCIENTIST

I have not attended the last URSI General Assembly, and I wish to remain/become an URSI Radioscientist in the 2006-2008 triennium. Subscription to *The Radio Science Bulletin* is included in the fee.

(please type or print in BLOCK LETTERS)

Name: Prof./Dr./Mr./Mrs./Ms. _____
Family Name *First Name* *Middle Initials*

Present job title: _____

Years of professional experience: _____

Professional affiliation: _____

I request that all information, including the bulletin, be sent to my ☐ home ☐ business address, i.e.:

Company name: _____

Department: _____

Street address: _____

City and postal / zip code: _____

Province / State: _____ Country: _____

Phone: _____ ext: _____ Fax: _____

E-mail: _____

Areas of interest (please tick)

☐ A Electromagnetic Metrology
☐ B Fields and Waves
☐ C Signals and Systems
☐ D Electronics and Photonics
☐ E Electromagnetic Noise & Interference
☐ F Wave Propagation & Remote Sensing
☐ G Ionospheric Radio and Propagation
☐ H Waves in Plasmas
☐ J Radio Astronomy
☐ K Electromagnetics in Biology & Medicine

The fee is 50 Euro.
(The URSI Board of Officers will consider waiving of the fee if the case is made to them in writing)

Method of payment: VISA / MASTERCARD (we do not accept cheques)

Credit Card No ☐☐☐☐ ☐☐☐☐ ☐☐☐☐ ☐☐☐☐ Exp. date: _____

Date: _____ Signed _____

Please return this signed form to:

The URSI Secretariat
c /o Ghent University / INTEC
Sint-Pietersnieuwstraat 41
B-9000 GENT, BELGIUM
fax (32) 9-264.42.88