

Review Draft. Comments welcome.

The Technical Basis for Spectrum Rights

Robert Matheson

Consultant to U.S. Department of Commerce/NTIA¹

and

Adele Morris

U.S Department of the Treasury²

June 23, 2007

The views expressed in this paper are not necessarily the views of the U.S Departments of Treasury or Commerce.

¹ Contact: egmatheson@att.net, (303) 497-3293

² Contact: adelemorris@verizon.net

TABLE OF CONTENTS

1.	ALLOCATION AND LICENSING OF SPECTRUM	34
1.1	Band Allocations.....	34
1.2	Spectrum Regulatory Authorities.....	45
1.3	Licenses and Assignments: An Introduction to Exclusive Spectrum Rights.....	56
2.	PHYSICS OF SPECTRUM USE	56
2.1	Electrospace Description of Spectrum.....	56
2.2	Frequency.....	8
2.3	Time.....	940
2.4	Physical (Geographical) Location.....	940
2.5	Direction of Propagation.....	1044
2.6	Dual-Space Spectrum Usage Models.....	1344
2.7	Exploiting More Electrospace Dimensions Expands Capacity.....	1546
3.	COMMAND AND CONTROL REGULATION OF THE AIRWAVES	1849
3.1	The Framework: Band Allocations.....	1849
3.2	Band Allocation Technology Examples.....	1920
3.2.1	Land Mobile Radio (LMR).....	1920
3.2.2	Point-to-point Fixed Microwave Services.....	2024
3.2.3	Radars.....	22
3.2.4	Broadcasting.....	2324
3.2.5	Summary of Command and Control Allocations.....	2425
3.3	Licensing.....	2526
3.3.1	License Allotments and Auctions.....	2526
3.3.2	Spectrum Scarcity and Electrospace Volumes.....	2627
3.4	Interference.....	2829
3.4.1	Co-channel interference.....	2829
3.4.2	Out-of-band interference.....	2930
3.4.3	Interference protection sets the boundaries of spectrum rights.....	3034
3.4.4	Interference disputes and tenant's rights.....	31
3.5	Receiver Performance Standards.....	3233
3.6	Duration of rights.....	3334
3.7	Command and control conclusions.....	3435
4.	ALTERNATIVES TO COMMAND AND CONTROL EXCLUSIVE RIGHTS	3536
4.1	Unlicensed spectrum.....	3637
4.2	Opportunistic Access.....	38
5.	AN ELECTROSPACE APPROACH TO FLEXIBLE SPECTRUM RIGHTS	40
5.1	Practical Electrospace Rights.....	41
5.2	Interference.....	4243
5.3	Setting E_{max} and E_0 Limits.....	43
5.4	The Problem of the Height Dimension.....	4647
5.5	The Probabilistic Nature of Propagation.....	4950
6.	DOCUMENTING AND ENFORCING SPECTRUM RIGHTS	5152
6.1	Rights Database.....	5152
6.2	Rights Protect Against Encroachment, Not Interference.....	52
6.3	Licensed rights and transmitters.....	53
6.4	Advantages and Disadvantages of Flexible-Use Rights.....	5554
7.	CONCLUSION	5755

INTRODUCTION

This paper explores the technical fundamentals of establishing rights to access spectrum, including institutional, scientific, and engineering considerations important to policymakers. It reviews how best to articulate rights to spectrum access such that rights holders can transfer, subdivide, aggregate, and protect their rights in a market-oriented policy environment. The paper examines current approaches to expressing rights to access spectrum, their advantages and disadvantages, and how they may lead to underutilization of spectrum resources. Section 2 outlines the basics of band allocations, frequency assignments, and spectrum licenses. This provides the foundation to examine the details of exclusive spectrum rights and assess the technical advantages and disadvantages of the current system with an eye toward identifying approaches that could work better. Section 3 describes the physics of radio signal propagation that underlie any spectrum applications and introduces the seven-dimension “electrospace” approach to describing radio signals and the rights to emit them. We argue that increased exploitation of these dimensions will be central in improving spectrum capacity.

Section 4 describes the traditional command and control approach to regulating airwaves that is the basis for most management of radio use in the United States and the rest of the world. Command and control regulations tailor specific frequency bands to specific types of systems, as shown by a description of the rules for bands allocated to mobile radio, microwave links, radar, and broadcasting. However, these specialized allocations for specific bands tend to be inflexible, and command and control management is often slow in providing spectrum that can rapidly be used by new technologies and applications. In addition, typical command and control management does not efficiently provide spectrum for certain uses. Section 5 briefly discusses several frequency management alternatives to command and control, including low-power commons and some opportunistic dynamic spectrum sharing techniques.

Section 6 describes how the electrospace approach can be the technical basis of flexible-use spectrum rights. We present a way to express the rights to use spectrum that is not tied to any specific service or technology. It would allow market forces to allocate spectrum such that new radio technologies and applications can be rapidly accommodated with minimal regulatory oversight. The rules needed to prevent interference while allowing substantial flexibility in use are relatively simple, though they are quite different from the rules that govern traditional command and control frequency use. For these and other reasons, it is likely that flexible-use frequency bands could fill some very important roles. However, it is also likely that many uses of spectrum will continue to be filled best by command and control and other types of spectrum management. Section 7 provides a summary and conclusions for this paper.

1. ALLOCATION AND LICENSING OF SPECTRUM

1.1 Band Allocations

In the U.S. and most other countries, regulatory authorities parcel out radio frequencies to users in a two-stage process. The first stage is *allocation*, whereby regulators divide the radio spectrum into frequency bands of differing rules. The rules in a band typically specify what services users can provide and the technical and operational parameters that apply to those

services. Figure 1 below shows the 2003 frequency allocation chart that applies to the United States. It shows that every frequency between 9 kHz and 300 GHz falls into a frequency band. The rules governing each band are specified in Section 47 of the Code of Federal Regulations and in the NTIA Manual of Regulations and Procedures for Federal Radio Frequency Management.

1.2 Spectrum Regulatory Authorities

Figure 1 shows that the Executive branch of the U.S. government, through the Department of Commerce’s National Telecommunications and Information Administration (NTIA), manages some bands for the benefit of federal agencies. Federal missions that require spectrum include law enforcement (e.g., the FBI, Coast Guard, and Secret Service), national infrastructure management (e.g., air traffic control by the Federal Aviation Administration, the Global Positioning System, and weather forecasting by the National Weather Service), and the safety of life and property (e.g., forest fire monitoring and patrolling of federal lands). A large amount of NTIA-managed spectrum is used by the military.

The Federal Communications Commission (FCC), an independent agency, manages the spectrum for all non-federal uses, including state and local government missions, broadcasting, private sector uses, and unlicensed applications (e.g. Wi-Fi, cordless phones). Finally, a few frequency bands are managed jointly by the FCC and NTIA.

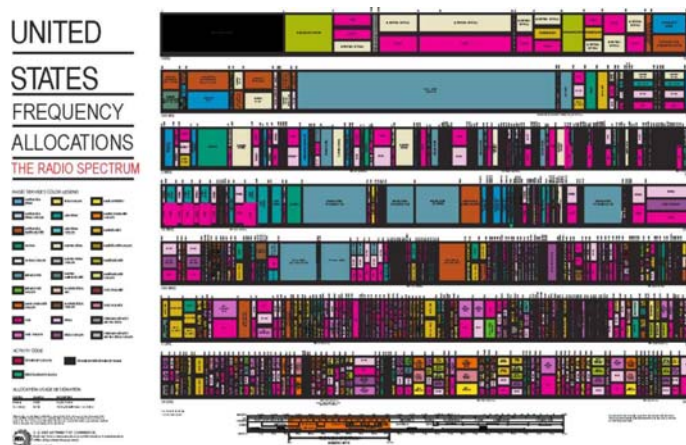


Figure 1. U.S. Frequency Allocation Chart

Some spectrum use, both federal and non-federal, is governed at least in part by international agreements. The U.S. makes two types of radio agreements with other countries. With immediate neighbors (chiefly Canada and Mexico) bilateral treaties describe how frequencies are to be shared in the areas adjacent to shared national borders. For example, the Canadian government must not issue spectrum access rights near the border with the U.S. that conflict with

rights on U.S. soil, and vice versa. More broadly, the International Telecommunications Union (ITU) of the United Nations coordinates the use of radio frequencies on a worldwide basis. The ITU standardizes band allocations within three different regions; the U.S. is in Region 2, which includes North and South America. The ITU regulates frequencies only in bands where the usage is inherently international. Such bands include many frequencies below 20 MHz (where ionospheric reflections carry signals around the world) and frequencies used by satellites.

1.3 Licenses and Assignments: An Introduction to Exclusive Spectrum Rights

The allocation table in Figure 1 does not specify which individuals may use a given frequency at a given location. That determination occurs in the second stage of governance, through *licenses* and *assignments*. In this paper, we will use the term “assignment” to denote the rights to access radio spectrum given to federal users and the term “license” to denote such rights for non-federal users. We will use the term “exclusive rights” to denote the rights contained in either a license or an assignment. Note that the term “exclusive rights” is importantly distinct from “exclusive use.” Indeed, spectrum that is licensed exclusively to a particular entity, such as a wireless communications provider, may be used by a great number of people. We use the term “exclusive rights” here to mean the legal right to exclude users who are not authorized by the spectrum rights holder. Exclusive rights stand in contrast to unlicensed spectrum, which can be used by anyone whose signals conform to the relevant rules for power levels and other technical parameters.

2. PHYSICS OF SPECTRUM USE

This section describes the physical properties of electromagnetic radio waves and explains the “electrospace” approach, a way of expressing the quantitative presence of those waves in the environment. We believe this approach is the best foundation for a flexible and technologically neutral articulation of exclusive spectrum rights.

Immutable laws of physics ~~govern~~affect the opportunities and limitations for all current and future spectrum uses, so this section seeks to make clear which constraints to spectrum use derive from radio physics and which constraints derive only from the existing regulatory structure or current technology. This distinction is critical to evaluating the current policy framework for spectrum management and identifying beneficial and practical options for improvement.

2.1 Electrospace Description of Spectrum

The term “spectrum” is used colloquially to mean several things, including a given frequency, a frequency band, or a set of rights to access a set of frequencies at a given time and location. In this paper, we use the alternative term “electrospace” to express the full potential for the extent

and coverage of radio signals through frequency, time, space, direction, and other dimensions.³ The electrospacetime describes the radio field strength at a given electrospacetime “location,” defined by a number N of independent electrospacetime dimensions. This means the electrospacetime represents an N-dimensional hyperspace.

The appropriate value of N depends on the number of characteristics of radio signals that current feasible receiver technologies can reasonably process independently from one another. **Table 1** shows seven key characteristics that today’s technology can usefully exploit, suggesting that at least for now, a seven-dimensional hyperspace could be sufficient for articulating a useful bundle of spectrum rights. We consider and reject two other radio signal characteristics, polarization and modulation, as potential independent dimensions suitable for subdividing electrospacetime volumes.

Table 1. Electrospacetime Dimensions

QUANTITY	UNITS	# OF DIMENSIONS
Frequency	kHz, MHz, or GHz	1
Time	seconds, hours, years	1
Spatial location	latitude, longitude, altitude	3
Direction-of-travel	azimuth, elevation angle	2

We will argue in this report that the electrospacetime approach is particularly useful for describing flexible, exhaustive, exclusive spectrum rights because it provides a straightforward and unique basis for specifying exact regions of spectrum (called “electrospacetime volumes”) to which an individual can hold rights of access. That is not to say that any particular rights holder would wish to subdivide his or her rights along every dimension. Rather, the practical value of enumerating these dimensions is to provide a flexible underpinning for an economically efficient, market-based allocation of spectrum across all its possible applications. Market forces and available technologies would determine the efficient level of partition of rights in any particular band or application.

In earlier literature, some researchers suggested a somewhat different set of electrospacetime dimensions than we do here. For example, many of the earlier investigators excluded the two “direction-of-travel” dimensions of azimuth and elevation because the technology at the time did not lend itself to useful exploitation of those propagation properties. It was difficult to construct

3. R. J. Matheson, “The electrospacetime model as a tool for spectrum management,” addendum to the Proceedings of the International Symposium on Advanced Radio Technologies, March 2003, NTIA Special Publication SP-03-401.

efficient directional receiving antennas for the much lower frequencies in use at the time, and adaptive directional receiving antenna systems and MIMO technology had not yet been invented.

In addition, some researchers include polarization and/or modulation as possible electrospatial dimensions, and we do not. Both signal properties are valuable tools in system design, but do not lend themselves to robustly distinct electrospatial volumes. Although polarization is surely a useful method of separating radio signals, we exclude it here because polarization involves only two possible orthogonal values, compared with an arbitrarily large number of different values for some of the other electrospatial dimensions. Moreover, whenever two polarizations are in use at a given location, it is likely that a single user will coordinate their use to ensure that the signals remain orthogonal to each other. Similarly, we have excluded modulation because any two differently-modulated signals must be distinguished from one another in a coordinated fashion. Of course, specific families of coded modulations exist that are orthogonal to other members of the same family (e.g., CDMA), but receivers cannot necessarily reject unknown modulations. The high degree of necessary cooperation between users of different polarities and modulations suggest that establishing regulatory boundaries between access rights across those dimensions would not be particularly useful, at least with current technology.

The three spatial dimensions identify the physical location of a hypothetical receiver at which place and time all the other dimensions (the field strength) are measured. We posit that a license for rights to an electrospatial volume would establish a numerical limit, e.g., E_0 in V/m, such that the licensee may not emit field strengths in excess of E_0 outside the licensed regions, perhaps with a de minimis exemption which we discuss in Section 6.6 below. A given signal is said to occupy an electrospatial “region” consisting of all locations in the 7-dimension hyperspace where field strength is greater than E_0 .

An “ideal receiver” can theoretically separate any two radio signals that differ in at least one of their electrospatial dimensions – even if they are present in the same geographic location. For example, two co-located radio receivers could function without interference if the signals were at different frequencies, or if the signals occurred at different times, or if the signals came from different directions. Likewise, ideal receivers can separate two otherwise identical radio signals (same frequency, operating time, and direction) if they are present at different locations.

The electrospatial model describes only transmitted radio signals and energy levels thereof. That means that an electrospatial model includes no consideration of receivers, and therefore no recognition of interference. Interference is a degradation of receiver performance caused by unwanted signals. It is a function of specific technologies, applications, operational and economic factors, and especially receiver capabilities. Interference may result from different signals occupying adjacent or intersecting electrospatial volumes, but it is not part of how we specify the electrospatial signals themselves. We discuss in Section 6 how an electrospatial rights approach allows markets to determine economically efficient levels of interference. Section 3.6 (dual-space spectrum usage models) describes a dual-space spectrum-use model, with a separate transmitter space and receiver space, that is much different from the electrospatial model. This dual-space model is appropriate where traditional command and control regulators need to consider total system performance, including interference and receiver performance.

2.2 Frequency

The frequency dimension of the electrospace has the standard meanings of the word, namely a description of the frequency or range of frequencies (bandwidth or pass band) at which field strength is being characterized. Frequencies can be divided over a wide range of increments, typically conforming to how regulators have divided particular services into channels. The term “frequency” is often shorthand for the range of frequencies (bandwidth) within which a system operates.

The physics of electromagnetic waves and the history of technology have affected how different frequencies are used today. The ionosphere often reflects frequencies less than 20 MHz back down to earth, which made these frequencies valuable for long distance communications in the eras before satellites and terrestrial/undersea optical cables were developed. In addition, vacuum tubes that formed the basis for early radio equipment worked best at the lower frequencies. Thus early radio devices operated mostly at lower frequencies, and their signals filled the bands below 20 MHz. After WWII, improved vacuum tubes (and later, transistors) raised the frequencies at which consumer devices could operate, but for many years the UHF TV bands (500-800 MHz) were at the top edge of usable consumer frequencies. As recently as the 1980's, 1 GHz was recognized as the upper frequency limit for consumer devices.

From the 1950's through the 1970's, special military, industrial, and scientific electronic devices operated at much higher frequencies (up to 10 GHz) to provide point-to-point microwave links and radars. These expensive and exotic electronic systems needed higher frequencies because the shorter wavelengths can be focused into tight beams of energy using proportionately smaller reflectors. Radar systems needed narrow beamwidths to send energy long distances and provide high angular resolution for detailed radar images. Point-to-point microwave systems take advantage of narrowly-focused radio beams and high-gain receiving antennas to efficiently pass communications between fixed relay stations in a chain of microwave links.

The focusing advantages of higher frequencies for high-gain, narrowbeam antennas are offset by the poorer performance of the omnidirectional antennas. Mobile and cellular systems that use omnidirectional antennas at higher frequencies typically adjust to this trade-off by building systems that use shorter-range radio links⁴

The commercial exploitation of higher frequency bands continues as technology improves. For example, the personal communication systems (PCS) in the frequency band 1850-1990 MHz are highly successful. Consumer Wi-Fi systems now operate at 5 to 6 GHz and consumer satellite TV receivers operate in the 12 GHz band. The development of cheap and powerful consumer electronics for the higher frequencies has greatly expanded the amount of electrospace that is

4. J. M. Vanderau, R. J. Matheson, E. J. Haakinson, “A technological rationale to use higher wireless frequencies” April 1998, NTIA Report 98-349.

suitable for consumer and other systems, opening new “frontiers” of usable higher frequency bands, as shown in Figure 2⁵

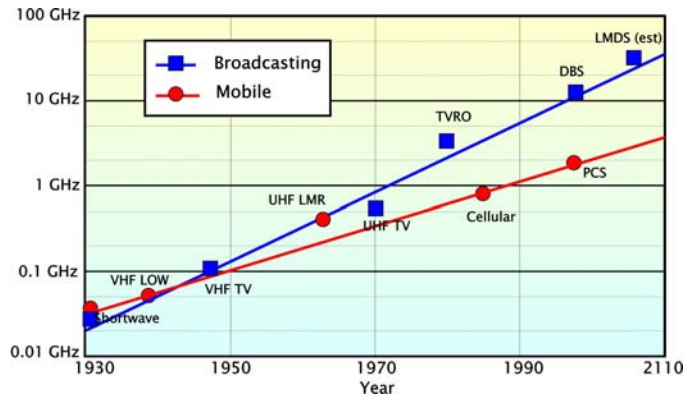


Figure 2. Frequency versus year-of-introduction for new consumer radio systems

2.3 Time

Time can be subdivided over a wide range of increments. Useful time divisions might include the several-year duration of a license, an agreement to allow a particular user to transmit regularly during the midnight-to-5 AM time block (when bandwidth would be inexpensively available to update computer files for the following day), or a one-time use during a 3-hour special events broadcast. On a much smaller time scale, a user could use a particular time slot on a TDMA system, to broadcast for a 2.5 millisecond time slot once every 20 milliseconds, or transmit data during the vertical blanking interval of an NTSC television signal 30 times every second. Many proposed cognitive radio systems presume that future radios will find and use frequencies that are temporarily unused in a specific time and place – possibly for free or by paying a fee. In its 2004 secondary market rules, the FCC encouraged the resale (or leasing) of unused spectrum by licensees, including temporal subdivisions.

2.4 Physical (Geographical) Location

The three spatial dimensions describe the physical locations where radio energy is present. Figure 3 shows an example of how a signal at 825 MHz may propagate over a geographic area, where the different colors show different signal levels (white being the strongest signal and pink the lowest). Although it would be more convenient for planning if the signal smoothly covered a circular area, Figure 3 shows that the world is typically more “messy.” In typical hilly terrain or within a city, many distant locations could have higher signal amplitudes than many closer locations.

⁵ ~~Error! Main Document Only.~~ Vanderau, p 21.

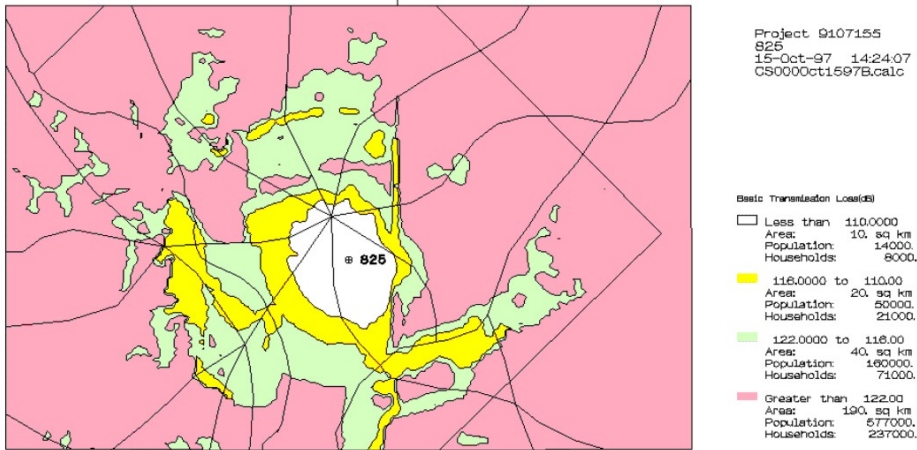


Figure 3. Propagation path loss at 825 MHz

This picture illustrates that the coverage area of a transmitted signal might include holes and drippy spills into outlying areas. In order to prevent excessive signal levels (larger than a limit E_0) outside the boundaries of a selected spatial region, users may have to greatly diminish signal amplitudes at useful locations within their spatial boundaries. Thus keeping a transmitter within specified spatial boundaries relies greatly on setting the right transmitter power, choosing a transmitter location, and using the appropriate directional transmitting antennas given the details of the terrain.

Along with longitude and latitude, the three spatial dimensions of an electrospacetime volume include height above the terrain, absolute altitude, or some other measure to indicate a vertical dimension. Figure 3 shows the predicted signal strength at about 6 feet above the level of the local terrain (possibly the typical height of a mobile antenna on a car or a handheld carried by a walking person). The area in white in Figure 3 would tend to increase in size and become more circular as altitude increases and the effect of ground-level obstacles diminish. Eventually the height reaches a point at which propagation is line-of-sight, and the signal strength decreases quite slowly with distance. Often blockage by the gross curvature of the Earth is the major limit on signal coverage in these cases.

2.5 Direction of Propagation

The direction of propagation of radio signals at a given location is another dimension of the electrospacetime that certain kinds of systems can exploit. However, only specialized applications can usefully exploit different directions of propagation because signal direction poses particular technical challenges. For example, although a transmitter induces signals to propagate mainly in the directions in which it is designed to emit, the signal may ultimately propagate in many different directions (as received in a single location) due to the scattering effects of the objects in its path.

Note that the direction of propagation of a signal is not necessarily determined by the angle at which a transmitting antenna points. At a given location behind a directional transmitting antenna, the strongest signal may come from the back side of the antenna or possibly from reflection off some object illuminated by the main beam. Therefore, the apparent direction of a signal at any given location does not necessarily correspond to the directional characteristics of the transmitting antenna. However, the directional characteristics of the transmitting antenna do affect the spatial locations where signals are strongest.

Receivers – including the pointing angle of receiving antennas – have no effect on the electrospace volume because they do not contribute to a signal’s intensity through the air. Rather, receivers with directional antennas can exploit the direction of propagation (for example, for point-to-point microwave systems, geosynchronous satellites, and radars) such that two electrospace volumes with otherwise identical characteristics may be distinguished from one another by the different directions of propagation of the signals.

In traditional radio systems, the only useful angle-of-arrival is usually the direct path between transmitter and receiver. For example, terrestrial microwave networks rely on point-to-point, free-space propagation between high-gain, narrow-beam antennas. These directional receiving antennas can very efficiently exploit the direction dimension by separating out individual signals at the same frequency from multiple microwave towers or geostationary satellite orbital slots. Figure 4 shows the directions of signals radiated from two omnidirectional transmitter sites. A receiver located within the small dashed circle would experience signals traveling in two different directions and could separate these signals using directional receiving antennas (even if the signals were at the same frequency). Notably, the *signal directions* at the receiving site allow this distinction, not the actual location of the transmitter sites. This technology is easily scalable by narrowing the beamwidth of the receiving antennas to separate out signals having smaller differences in direction at the receiving site.

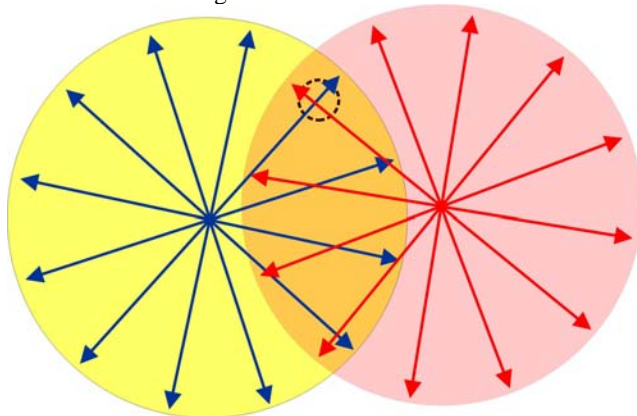


Figure 4. Directional antennas separate signals with different directions of propagation

In many applications, including most mobile and cellular systems, the directional attributes of the radio signals vary over time. Such applications cannot exploit the direction of arrival using simple directional receiving antennas. However, cellular/PCS systems are beginning to use adaptive receiving antenna systems that continuously adjust the directional characteristics of the receiving antennas to track the directional characteristics of the received signals. Such systems can substantially increase the ability of base stations to re-use frequencies without increasing interference.

In the right environment, a transmitter can generate multiple independent signals (at the same frequency) and strategically scatter them such that they come from different directions at the receiver. In that case the receiver can use multiple directional receiving antennas to receive multiple independent signals from the transmitter site at a single frequency. Figure 5 shows how this could be done.

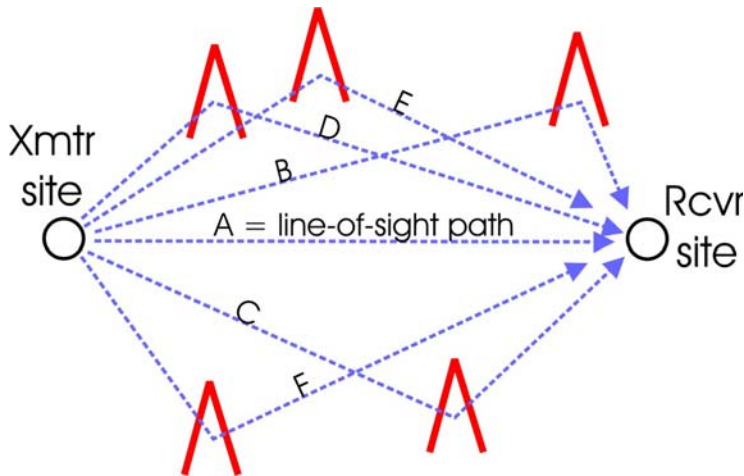


Figure 5. Multiple independent signals by strategic scattering

In a standard line-of-sight system, only Path A (the direct path between the transmitter and receiver) is normally used. However, if the transmitter site transmitted a higher power directional beam (at the same frequency) along Path B towards a selected scattering object (shown by the inverted “V”) that was mutually line-of-sight to both the transmitter and the receiver, a usable amount of signal would be scattered from the object to the receiver. Part of this signal would arrive at the receiver site coming from the direction of the scattering object. The receiver could use a high-gain receiving antenna that would pick up only the Path B scattered signal, giving a Path B signal that is independent of the original Path A signal.

Similarly, the transmitting and receiving sites could use additional paired directional antennas to bounce signals off other objects to generate additional independent paths C, D, E, and F. The scattering objects must be mutually line-of-sight to the transmitter and receiver, and they must all be at different angles as seen by the transmitter and receiver.

The direction of propagation is likely to be more important in the future than it is now. For example, advanced technologies, including adaptive antennas and Multiple Input Multiple Output (MIMO) systems exploit the angle of arrival to allow two or more times greater communication capacity and increased range with no greater transmitter power than conventional systems. Recently developed MIMO technology exploits multipath reflections and multiple transmitting and receiving antennas to generate independent transmission channels, somewhat like the multiple directional beams with different apparent angles-of-arrival shown in Figure 5. However, instead of using a combination of directional antenna beams to produce multiple independent paths, MIMO uses multiple omnidirectional receiving and transmitting antennas, whose signals are all bounced off various random objects, causing the reception of multiple signals with slightly different multipath components. Under certain (fairly common) conditions, the multipath signals received on the multiple receiving antennas can be mathematically processed to separate the independent signals that were transmitted by a set of multiple transmitting antennas, synthesizing multiple independent radio channels at the same frequency between a pair of transmitter and receiver sites. MIMO technology is potentially much cheaper and requires no painstaking aiming of directional antennas. The recent ~~(still under development)~~ IEEE 802.11n Wi-Fi standard uses MIMO techniques to substantially increase the distance and bits/s of data that can be transferred, compared to single-path Wi-Fi techniques.

This significant and growing potential to create independent, distinguishable signals at the same frequency in the same location supports our contention that direction of propagation is a useful electrospatial dimension.

2.6 Dual-Space Spectrum Usage Models

The electrospatial model of spectrum usage described in the previous sections is a model that recognizes only the existence of radio signals and ignores any aspects of receivers. We advocate this specific model as the technical basis for flexible-use spectrum rights because it describes the aspects of the physical world that can be measured (radio waves) and which are directly responsible for providing radio services or causing interference. We exclude receivers in the electrospatial model because they do not affect the radio wave environment, cannot cause malfunctions of other receivers or transmitters, and incorporating them unnecessarily complicates ascribing responsibility for interference. We argue that the electrospatial model is the best way to express the flexible-use rights discussed in Section 6.

The electrospatial model contrasts sharply with the spectrum rights approach employed by most command and control spectrum managers for the past century, the “dual space model.” Traditional spectrum regulations implicitly contain two interacting “spaces” having the same nominal seven dimensions as the electrospatial model.⁶ One of these spaces is known as the “receiver-denied-space” or “receiver-space;” the other is known as the “transmitter-denied-space” or “transmitter-space.” The transmitter-denied-space consists of the seven-dimension region where

6. D. L. Ewing, L. E. Berry, “Metrics for spectrum-space usage,” OT Report 73-24, November 1973

signals from transmitters are present. A given location is “occupied” if the presence of signals at that location (geographical, temporal, frequency, and direction) would deny that space to a new receiver belonging to another system, because the new receiver would get interference from the existing transmitted signals.

Receiver-denied-space is the “electrospace” that is “occupied” by receivers. Receiver-denied-space is considered occupied at all locations where a new transmitter cannot be licensed, since it would cause interference to the existing receivers. Therefore, the existing receivers deny the use of that spectrum-space to new transmitters. Since receivers don’t radiate signals, however, the receiver-space is occupied by a *regulatory obligation* to protect the interference-free operation of licensed receivers.

At first glance, the transmitter-space may seem almost identical to the electrospace. However, they are different. The transmitter-space and the receiver-space do not contain signal-strength numbers, like the electrospace does. Instead, the transmitter-space and receiver-space contain occupancy states, whose values are either “yes” or “no,” i.e., “occupied” or “not occupied.” If the transmitter-space is occupied at a certain seven-dimensional location, it cannot be occupied by a receiver belonging to new system. If the receiver-space is occupied at a certain location, a transmitter for a new system cannot be licensed there. Assuming the use of omnidirectional antennas (true only for some types of systems), transmitter-space occupancy will tend to be centered on transmitter locations, while receiver-space occupancy will tend to be centered on receiver locations.

The dual-space concept works only in bands that are allocated to support a single type of specific system and service having known transmitter and receiver characteristics. These models allow spectrum managers to provide licenses that will operate properly and without interference, having taken into account the effects to and from all other systems that have already been licensed in a given band. The spaces are crucially dependent on the assumed characteristics of the transmitters and receivers that are used in the particular band, since the decisions about whether given locations are occupied will depend on the specific transmitter and receiver characteristics.

Transmitter-space is defined as “occupied” at all locations where a receiver for a new system cannot be added (because the new receiver would experience interference from one or more existing transmitters). Transmitter-space must be analyzed before a new receiver can be licensed. Such analyses could be done using terrain-based propagation models, using existing transmitter characteristics and the characteristics of a *reference receiver*. The characteristics of the reference receiver are typically the characteristics of the receiver that is recommended or specified for operation in that specific allocated frequency band and service. For, example, if the front end overload characteristics of the reference receiver prevented proper receiver operation whenever the receiver is operated within 200 meters of the transmitter over a 100-MHz range of frequencies, the transmitter-space will be counted as occupied over the 100-MHz range for geographical locations within 200 meters of the transmitter location. On the other hand, if the proposed new receiver used a directional receiving antenna that prevented reception of the signal in most directions, the occupied transmitter-space would need to be re-calculated to show the

lack of transmitter-space occupancy in locations where the new receiver antenna pattern would adequately reject transmitted signals.

Receiver-space is defined as “occupied” at all locations where a transmitter for a new system cannot be added because it would cause interference to existing receivers. Receiver-space must be analyzed before a new transmitter can be licensed.

The process of awarding a new license for a new radio system would theoretically involve computing the existing receiver-space and the existing transmitter-space to be sure that receiver-space is not already occupied at the location selected for the transmitter and that the transmitter-space is not occupied at the location selected for the receiver.

Various allocated bands could have considerably different versions of the receiver-space and transmitter-space, depending on the details of services provided in that band allocation. For example, point-to-point microwave bands would probably use an 8-dimension space that includes an additional dimension for polarization. Mobile radio bands would use five dimensions, including neither direction nor polarization. The inclusion of two quite different sets of equipment in a typical mobile band (i.e., the fixed-location base stations and the mobile units with constantly changing locations) means that quite different assumptions will need to be made about how mobile units are protected from interference versus how base stations are protected. Broadcasting services are sufficiently different from other services that the dual-space analysis includes only the receiver-space.

The dual-space model is especially applicable to command and control frequency bands, where each band contains a uniform population of transmitters and receivers, and where there is an obligation to ensure that receivers are protected from interference. The regulatory obligation to protect receivers means that the electrospace described in Section 3 must be augmented to include information on receivers. It also means that the process of deciding who is responsible for correcting interference problems is conceptually much more complex, especially if all areas of technical performance of the pertinent radio devices are not specifically enumerated. The extra complexity is associated with the ambiguity of whether interference is caused by a faulty transmitter or by a faulty victim receiver, as well as other factors such as which system was licensed first at a specific location. There may be especial ambiguity about whether wide-area mobile systems can be considered to be licensed at any specific locations.

2.7 Exploiting More Electrospace Dimensions Expands Capacity

The amount of information carried by wireless systems is growing rapidly across the globe. Ten to twenty years ago, some were concerned that usable spectrum would soon be fully exploited, precipitating a major communications crisis. Today, however, analysts view spectrum capacity, the ability of spectrum to carry communications, as more elastic.

By itself, raw spectrum has no ability to carry communications. Spectrum capacity appears only when the proper equipment (transmitters and receivers) is installed and operated. A great variety of equipment exists, with an accordingly wide range of spectrum capacity. For example, a 6-

MHz broadcast TV channel can carry one standard-definition TV (SDTV) program using old analog NTSC technology, or five SDTV programs using digital TV with MPEG-2 compression techniques, or possibly ten SDTV programs using future MPEG-4 or similar compression techniques. Thus, the number of TV programs that can be carried in a single 6-MHz-wide channel depends on the technology and infrastructure deployed.

The TV example shows how spectrum capacity can be a function of advanced image-compression processing. Many other spectrum capacity improvements result from more intensive exploitation of electrospace dimensions. For example, short-range systems (in contrast to longer-range systems) allow the same frequency to be reused more often on a geographical basis. Thus, smaller cells in PCS or cellular systems, or short-range Wi-Fi systems increase spectrum capacity by geographically partitioning the electrospace. Trunked radio systems partition a channel by time across many users, allowing the system to carry much greater amounts of traffic on each frequency. The number of frequencies available to carry the rapidly-growing amounts of consumer wireless traffic has increased almost five times over the last decade or two, as the upper frequency limit for economical consumer systems has increased from 1 GHz to about 5 GHz. And as discussed in Section 3.5, the electrospace directional dimension is now involved much more intensely in systems using MIMO technology to generate multiple paths at the same frequency between transmitter and receiver (e. g., Wi-Fi standard 802.11n).

Different systems exploit different dimensions of the electrospace, and the approach a user employs in any given context depends on the available technology and market factors. Since every electrospace dimension can allow additional independent information paths between the transmitter and receiver, the use of more complex systems can in theory greatly multiply the amount of traffic carried in a given band. For example, a new system that might combine a MIMO technique that increases direction-of-arrival capacity by a factor of K , a small-cell architecture that improves spatial reuse by L times, and a modulation/compression technology that increases the traffic carried in a given bandwidth by M times could produce a total improvement in system capacity by a factor, $F = K \times L \times M$.

Matheson[1994] discussed many of these techniques to improve spectrum capacity. An excerpt from that paper validates a prognosis of increasing exploitation of electrospace dimensions to expand spectrum capacity:⁷

“SUMMARY OF SPECTRUM CAPACITY FACTORS

The previous sections have described some of the factors that are expected to produce additional spectrum or spectrum capacity. There are more factors that could generate additional spectrum capacity, such as trunking technologies, more precise modeling in frequency management decisions, the use of active interference avoidance in low-power bands, and market-based redistribution of under-used frequencies. In this section,

7. R.J. Matheson, “Spectrum stretching: adjusting to an age of plenty,” IVHS Journal, 1994, Vol. 1(4), pp. 397-407.

however, I will consider only the cumulative effect of spectrum capacity factors discussed earlier:

- a. Fiber optical spectrum reclamation 1.2
- b. Federal/military frequency reclamation 1.1
- c. Availability of higher frequencies 5
- d. Frequency reuse/short-range systems 25
- e. Digital compression techniques 3

Total increase in spectrum capacity = 1.2 x 1.1 x 5 x 25 x 3 = 495

The product of these factors gives an overall increase in spectrum capacity of 495. This means that within the next 10-15 years these factors are expected to give us the equivalent of 495 times the spectrum capacity that we have at present using traditional technologies. It is worth noting that the factors that make more frequencies available (a,b,c) give an increase of 6.6, while the factors making better use of frequencies (d,e) give an increase of 75. The reclamation-based factors (a,b) increase the spectrum by 1.3, while the technology-based factors (c,d,e) increase spectrum capacity by 375. This suggests that technology improvement has much more payoff than a rigorous reclamation of frequencies from lower-priority users.”

An important conclusion from the above excerpt is that improved technology and investment can greatly increase the amount of available spectrum capacity (somewhere around 375 - 495 times, according to the above summary). In the 13 years since the paper was written, many of the predicted improvements have been realized or are recognizably progressing. If Matheson[1994] were rewritten today, it would include somewhat different factors and much of the improvement in the old factors would be part of the baseline from which we are starting today. New factors might include MIMO techniques, which had not yet been developed 10 years ago, which might give another factor of 2-3. Dynamic sharing of unused frequencies and wireless mesh networks were not yet being widely discussed then, either, and might be included today.

However, if new technologies are to incorporate these improvements, the radio regulatory framework must allow their rapid and flexible deployment. History suggests that technological improvement flourishes in the most flexible regulatory environments. Over the past 10 to 15 years, a disproportionate share of improvements in spectrum capacity occurred in the unlicensed bands (via short-path systems – especially Wi-Fi – and higher frequencies), the cellular/PCS bands (via short-path systems, compression, and fiber optic reclamation), and the satellite broadcasting (TV and audio) bands (via compression, directional discrimination, and higher frequencies). The traditional command and control bands, which we discuss in the following section, did not typically experience similar improvements. One possible exception is the broadcast TV digital transition which is still underway, since it will deploy important spectrum capacity improvements through digital image compression. Licensed usage in many of the huge federal and non-federal point-to-point microwave and MMDS bands actually dropped, as the

traditional services provided by these bands were replaced by more modern alternatives, especially optical fiber.⁸

Suppose we assume that over the next 10 to 15 years, unknown new systems will be able to economically deploy improvements in the spectrum capacity similar to the past decade, and suppose our ability to foresee these new systems is no better than a savvy observer 15 years ago to foresee GPS/camera/video/MP3/SMS/cellphone systems. The following sections explore what we should do now to ensure that the spectrum management system does not forestall the enormous improvements otherwise made feasible by new technology.

3. COMMAND AND CONTROL REGULATION OF THE AIRWAVES

3.1 The Framework: Band Allocations

The United States, through the FCC and NTIA, currently allocates bundles of exclusive spectrum rights (licenses or assignments) primarily through command and control regulation. In some cases, licensees can allow others to access their licensed rights. For example, cell phone users are allowed to emit signals on the frequencies controlled by the cell phone system operators. Most fundamentally, the regulatory authorities determine which spectrum is accessed by which types of radio systems. This section describes how the current approach works, with special attention to the advantages and disadvantages of command and control regulations for specific kinds of spectrum-dependent applications.

The FCC and NTIA have reserved specific frequency bands for specific services, covering all currently useful frequencies. The range of frequencies they regulate lies between 9 kHz and 275 GHz (the “radio frequency range”), although currently no devices operate above about 100 GHz. Optical devices (with frequencies 10,000 times higher than 100 GHz) also use electromagnetic waves, but at this time these are specifically exempted from regulation as radio devices.

Within the radio frequency range, every frequency falls into exactly one allocated band. In general, a single set of rules applies to each frequency in a band, although the U.S. allocation table annotates some individual frequencies with special footnoted rules. The broadly applicable rules within a frequency band dictate what services the user can offer, by whom, using what signal characteristics. For many years, these rules were even more specific, for example specifying exactly which type of mobile radio user (such as forest products, movie production, or ground transportation) could use a given frequency. Over the past few decades, regulators have eased band allocation rules somewhat, but important strictures remain. For example, limits remain on the ability of users to change services they offer and to transfer, subdivide, aggregate or otherwise flexibly manage their spectrum rights.

Let us consider the regulation of a single command and control frequency band. The most important limitation in a frequency band regulation is the single type of radio system allowed to access spectrum in the band, along with the highly specific rules about how the systems must be

8. R. J. Matheson, “Spectrum usage for the fixed services,” March 2000. NTIA Report 00-378

configured. In a mobile radio band, for example, a spectrum license usually specifies the exact frequency and bandwidth of operation; the location, height, and gain pattern of the base station transmitting antenna; the maximum transmitter power and modulation, the function of the communications link, and more. In some bands, a user may face a menu of options that can allow better adaptation across rural or urban operation, heavy traffic versus low traffic, long range versus short range, and other conditions. The regulatory agency typically provides well-engineered system designs for each allocated band, with the expectation that a radio system built with standardized designs will function without causing or experiencing interference. These rules also set typical or worst-case geographical separation distances. Channelization bandwidths describe the width of frequency bands that regulators supply to each user. Emission masks describe how much energy users can leak into adjacent frequencies. In most cases, industry is well prepared to supply electronic equipment that specifically matches all of the technical requirements specified in the band allocation rules.

Regulators have occasionally adjusted rules for specific types of service as technology, experience, and market demands evolve. Allocation rules often change at the instigation of an industry/user committee, possibly via a series of public hearings. These hearings typically start with an FCC notice of inquiry (NOI), followed by a notice of proposed rule-making (NPRM), and end with a final rule that incorporates a modified set of band allocation rules. Such proceedings can require a year or more, depending on whether opponents attempt to delay or modify the changes (possibly for competitive/business reasons or for fear of interference to existing systems).

3.2 Band Allocation Technology Examples

The band allocation process has been used to develop specific rules for a wide range of radio services. This section describes rules for a few examples, with the intention of illustrating the range of technical problems that such rules need to solve.

3.2.1 Land Mobile Radio (LMR).

LMR transmits signals between base stations and users in cars or walking with handheld radios. The range of the service of most systems is typically 20 miles, but higher antennas and more transmitter power can greatly extend it. LMR systems generally use omnidirectional antennas, which are best suited to a continually changing user location. The communications channel is designed mostly to carry speech, so it can be relatively narrowband. Analog LMR radio channels are as narrow as 15-25 kHz, while LMR radios that use digital compression can be as narrow as 5-12 kHz. Since a typical LMR radio user might use the radio channel for only a few minutes or tens of minutes per day, the efficient use of a radio channel means that many LMR systems may be licensed to share a single channel among multiple independent users. Strong “courtesy” conventions have developed that require a user to listen for current traffic among other users and wait until the current conversation has finished before initiating his or her own conversation.

Special challenges for licensing LMR systems arise from the wide range of potential transmitter locations, including hilltops and rooftops (where their radios have a greatly increased range),

inside basements (where signals will have an especially decreased range), close-in to the transmitters, and distant locations where reception is near its limits. These factors, along with the random timing of usage, challenge service providers' ability to plan for specific guaranteed levels of performance. But because of the high demand for LMR service, regulators have channeled the LMR as narrowly as possible, reducing channels from 50 kHz historically to 12.5 kHz today (with prospects of soon dropping to 6.25 kHz).

More recently, trunked radio technology (where multiple channels are available at a given base station, and a user is automatically switched to an arbitrary unused channel whenever service is needed) has begun to substantially replace single-channel radio technology. Thus, the FCC set new LMR rules that allowed (and encouraged) trunked systems to be assembled from multiple single channels. Unfortunately, service providers did not always manage the transition to trunked radio systems correctly, and newly-built trunked systems in the 800 MHz band caused interference to other LMR radios in the band. LMR services are also facing major competition from cellular/PCS systems, which will probably result in the demise of many LMR systems.

Technology-wise, cellular/PCS systems are essentially trunked LMR systems that have evolved to handle the very dense user environment in large metropolitan areas. By using smaller cells and more flexible modulation formats, cellular/PCS systems provide a wider range of services and much more total traffic per MHz than traditional LMR services. The smaller cells allow greater frequency reuse and wider data bandwidths for a given transmitter power, but they require much more (~~expensive~~~~costly~~) infrastructure. In addition, cellular/PCS systems provide simultaneous voice channels in both directions (full-duplex), whereas LMR usually provides a one-direction, push-to-talk voice channel.

Several features characterize a typical LMR system. Small businesses and local government agencies, like trucking companies, taxi services, and local police and sheriff departments, often deploy LMR systems for use by personnel who have little technical (electronics) training. The large numbers of mobile and portable terminals means that the radios must be kept small, light, inexpensive, and durable. LMR systems can take advantage of lower frequencies to improve the potential distance over which the radios can transmit useful signals with omnidirectional antennas. The current band plan includes many LMR bands, ranging from 30-50 MHz to higher than 900 MHz.

In summary, the LMR bands use extensive division of the electrospatial time and frequency dimensions, moderate division of the spatial dimension, and little division of the directional dimension.

3.2.2 Point-to-point Fixed Microwave Services

The point-to-point microwave services are in many ways orthogonal to LMR service. Typical microwave services use a series of fixed base stations, positioned in a multi-site chain, so that they can relay wideband data from one end of the chain to the other, generally carrying traffic in both directions. Since the stations are all fixed in location and communicate only with the two adjacent stations (line-of-sight), they can use highly directional antennas. Microwave systems

typically carry continuous wideband signals that may combine a large number of independent narrower bandwidth signals.

The factors that have characterized point-to-point microwave links include use of higher frequencies (1.7-24 GHz) because those frequencies are better suited to high-gain directional antennas and bandwidth is less scarce. Also, many users of microwave links were large corporations and federal agencies who had access to the necessary high tech equipment necessary. Generally the equipment is located in buildings, where calibration and maintenance is easier. Because of the narrow antenna beamwidth (e.g., 1-5 degrees), microwave frequencies can be reused many times in a single geographical area. Users have also successfully packed in multiple data bits per sample of data with multi-level modulations. These complex systems have the advantage of being fixed in location and time, so that a single, well-engineered system can typically remain in place for years or decades, operating continuously.

The major technical challenges in allocation rules for point-to-point microwave bands concern how well users can control the directional properties of radio waves. State-of-the-art propagation models with accurate terrain and obstruction data, along with certain details on climate and precipitation data, are vital for adequate planning of microwave systems. Directional antenna technology (especially second-order effects like sidelobe and backlobe responses) is highly important. Rules may require antennas with narrower beamwidths and lower sidelobe/backlobe in metropolitan areas where the bands are relatively crowded and higher frequency reuse is valuable.

Three technological changes raise questions about current rules for point-to-point microwave. The first is the development of optical fiber as the chief conduit for most broadband data, largely replacing point-to-point microwave in the role for which point-to-point microwave frequency bands were so generously allocated in the past. The more-than-adequate availability of microwave spectrum is underlined by recent FCC rule changes in the 10 GHz band and proposed changes in the 11-GHz band, allowing the use of smaller antennas (one of the rare recent instances of regulators reducing spectrum efficiency).⁹ The second technology development is the slow but inexorable improvement in high-frequency electronic devices, which moves inexpensive consumer electronics into continually-higher frequency bands (possibly up to a 5-GHz upper limit today). Thus, microwave bands that previously required exotic and expensive ultra-professional-grade electronics (limiting their usefulness to a narrow class of applications) are now becoming attractive for many other services. Point-to-point bands that were previously free of competition with consumer services may unexpectedly find themselves considered as the new “beachfront” properties for consumer services. Finally, geostationary satellite services have been added to many terrestrial microwave bands. Although these terrestrial and satellite services share quite well with each other, many newer terrestrial services may not share well with the satellite services. This may substantially limit the reallocation possibilities for terrestrial microwave bands.

⁹ FCC 07-38. WT Docket No. 07-54. NPRM for Amendment of Part 101 of the Commission’s rules to modify antenna requirements for the 10.7-11.7 GHz band. March 27, 2007

In summary, point-to-point microwave bands heavily exploit electrospace use on the basis of the directional dimensions and the associated geographical dimensions. In addition, historically microwave services mainly used the higher frequencies (above 3 GHz) which were uniquely suited to microwave point-to-point system design and at which most other uses were neither technically possible nor economically viable. The use of the time dimension is essentially ignored for the frequency management of these systems, although mobile/portable electronic news gathering (ENG) microwave system frequencies are often timeshared among television stations for local news coverage.

3.2.3 Radars

Radar systems are non-communications (sensing) applications for various navigation and military purposes. Radars in mobile applications (especially ships and airplanes) detect the vehicle location relative to terrain, as well as other vehicles, docks, runways, and other relevant objects. The first radars used very high power pulsed signals (possibly a one million watt pulse lasting one microsecond in duration and repeated 1000 times each second) with a very high gain (30-40 dB) and highly directional antenna. Radar antennas rotate in direction (typically 360 degrees azimuthally every 5-10 seconds). Any radio-reflective objects in the antenna main beam reflect (echo) some of the pulse energy back towards the radar transmitting antenna (which switches to receiving mode immediately after transmitting the pulse). Observers know the distance from the radar to the object by carefully measuring the exact timing of the received echo. The amount of energy received in the echo gives some indication of the potential size of the object. With the rotation, the system completes a picture of the location of all objects surrounding the radar transmitter within a range of 50-250 miles every 5-10 seconds.

Early radars generated the short powerful pulses using a “magnetron,” which tended to radiate energy across a wider-than-necessary band of frequencies. These relatively dirty signal sources caused nearby radars to interfere with each other, initially requiring a fairly large frequency separation between radars. More recently, the magnetrons and other transmitting devices have become cleaner, and radar receivers using pulse blankers and correlators eliminate much of the radar pulse trains that have pulse repetition rates (PRRs) that are different from the PRR of the local radar. Some recent radars use longer duration coded pulses with a series of phase reversals. These long-pulse radars often use solid-state power amplifiers (instead of magnetrons), have much cleaner spectral outputs, provide additional ways to reject interfering radar pulses, and decrease the potential of high-power radar interference to other radio systems.

Some radar applications face complicated military and security conditions. For example, military radars must work in an environment of jamming and deception, a mission requirement that dominates over concern for greater spectrum efficiency. In particular, military radars may need to change frequencies, waveforms, signal processing, and beam-scanning rapidly and adaptively. This can help radars escape jamming and spoofing, detect “stealth” targets, and avoid incoming missiles homing on the radar signal. These mission requirements may limit how technically efficient these applications can be.

Indeed, radar bands may exhibit minimal observed transmission energy relative to conventional radio systems. The very strong radar transmitter signal is typically present for only one part in 100,000 of the time at a given location. The very weak radar echoes (that constitute the important part of the radar signal) are present a much larger portion of the time, but they are usually too small to be seen by normal radio receivers. Therefore, radar bands may seem unduly empty even when fully utilized. Possibly partly because of their “empty” appearance, regulators have selected some radar bands near 5 GHz to be shared with unlicensed wideband wireless data devices, using a sharing technique called “dynamic frequency selection” (DFS). DFS requires the unlicensed wireless device to continually search for radars and to change to another frequency if a radar is detected.

Radars use roughly the same frequencies as point-to-point microwave systems for the same reasons: bandwidth availability and the need for highly directional antennae. Higher frequencies are especially important for radars integrated into aircraft, missiles, and other smaller vehicles. Therefore, radar bands and point-to-point microwave bands are often located adjacent to each other. Radars cannot generally operate in the same bands with point-to-point microwave links and satellite systems because the high power radar pulses can overload microwave receivers.

The degree to which reform of radar bands is feasible is uncertain because operational demands by the military are uncertain. On one hand, many radar bands are relatively empty (in peacetime, at least). Further, global positioning system (GPS) technology provides much cheaper, more accurate positioning information than radars, potentially replacing older radars for many navigation and location uses. New coded-pulse radar technologies could make it possible to squeeze multiple radars into a much smaller set of frequencies than they now use. On the other hand, a large inventory of very important military radars may operate most effectively when they can change frequencies over a wide range on a pulse-by-pulse basis.

In summary, radars make extensive use of the electrospatial dimensions of direction and time, with less exploitation of frequency and location. Future non-military radar bands could possibly use a combination of GPS and coded-pulse technologies to operate in much less spectrum than current systems. However, even many nominally non-military radars have some military or security functions, and military radars often need to use spectrum tactically in relatively non-efficient ways.

3.2.4 Broadcasting

Broadcasting is probably one of the earliest systematic uses of radio. A major goal of broadcasters is to deliver a usable signal to a large number of listeners/viewers in whatever local situation they are found. Broadcasting is different from other radio services because it operates only in one direction, from the studio to the listener. Broadcasting today is tightly regulated by a number of technical policies, including specific technical broadcast standards that allow free reception ~~free~~ by millions of privately-owned inexpensive receivers. In addition, many non-technical constraints apply, such as public service announcements and warnings, political campaigning rules, station ownership limitations, and obscenity censorship.

The scheduled conversion from older analog TV to digital TV and possibly to some form of digital (FM) radio has prompted a flurry of arguments and delays, for example to include protection of the intellectual property (IP) rights of the broadcast content. IP was a smaller problem with earlier analog formats because the quality of the program material degraded substantially with each additional stage of re-recording. With digital content and recording, however, each copy is perfect, and copies recorded off-the-air are essentially identical to the original programming. This raises issues like the performance of copy-protection systems and whether new rules allow *any* digital material to be sent to listeners – not just music or video. It is also uncertain how much freedom the viewer will have to adjust the viewing experience, for example to excise commercials, send the programming from the set top box via Wi-Fi to the home-theater digital display, and to record the program on a video iPod

Two facts bear most critically on the future of broadcast spectrum. First, most viewers get TV via cable or satellite, with a minority receiving TV via a broadcast signal. Second, with the switch to DTV, the selected 8-VSB broadcast video modulation technique will not (barring any major technical breakthroughs) provide usable signals in many locations where analog TV currently works – including most mobile or portable receivers. Thus, one would expect the number of over-the-air viewers to drop even further after the DTV conversion. It is true that the DTV conversion will reduce the number of TV channels from 59 to 43, but each DTV channel could transmit as many as 10 standard definition TV (SDTV) channels. Therefore, the switch to digital TV could actually *increase* the number of broadcast TV programs from 59 to 430 (ten SDTV programs per channel). Even with this potentially huge increase in broadcast programs, few viewers may watch.

Public policy has traditionally reserved a special place for broadcasting. Most of the capital investment in broadcasting (especially TV) is in the receivers purchased by the public, so some argue that it would be unwise to allow broadcasters to “orphan” all of this investment. Thus regulators have restricted options to broadcast signals to certain formats. In particular, regulators may have locked TV broadcasting into a single somewhat deficient 8-VSB broadcasting standard that dooms the industry. For example, TV broadcasters have not sought ways to transmit DTV to cell phones or any other mobile appliance. Broadcast in 8-VSB competes with cable, DSL, optical fiber, and satellite for fixed viewers, and non-fixed users cannot successfully receive the signal. Broadcasters are not considering providing other services with its 20-MB/s digital service. The broadcast link won’t work well enough for most applications, and broadcasters cannot legally transmit non-8-VSB signals.

In summary, broadcasting is a favored but increasingly less-consequential use of the radio spectrum. The difficult over-the-air DTV broadcasting standard, the confusion and inconvenience of IP-protection rules, and strong digital-program transmission competitors all point to an uncertain future for DTV broadcasting. Similarly, iPod and other devices will substantially affect the broadcast market. The broadcasting industry moderately exploits the electrospace dimensions of frequency and location, with little division of access by time or direction.

3.2.5 Summary of Command and Control Allocations

The rules that apply within each of the band allocations described above result from the particular technical components required to provide those respective services. These band allocations provide distinct optimized radio services and allow vastly different services to operate in appropriate electrospacetime volumes. This includes a wide range of frequencies – from the lowest of ocean-penetrating 20 kHz submarine communications to the 60 GHz band where huge bandwidths are available over short line-of-sight paths (where frequency re-use is aided because signals are strongly absorbed by the oxygen in the atmosphere). For each allocated band, the command and control approach allows carefully tailored rules and standards for well-optimized systems with good spectrum efficiency and moderate cost. The set of regulations and standards for many band allocations constitute a virtual blueprint for the design of a wide variety of complex radio systems. In many ways, the command and control management of frequencies has served spectrum users very well.

With new technologies and spectrum demands, the command and control approach could give way to a future policy portfolio that allocates most spectrum under flexible-use property rights rules (Section 7), rules that allow various types of opportunistic (dynamic) sharing (Section 6), and hybrid approaches that capture some features of licensed bands and some features of unlicensed bands. Each of these approaches provides an advantage over traditional radio-system specific rules. Unlicensed bands allow low power transmissions from very-short-range devices without a license. This is particularly valuable for applications in which the transaction costs of licensing users would far exceed the value of the small quantity of electrospacetime that they consume. The flexible-use spectrum rights could accommodate competitive new services and allow more efficient market allocations of rights across exclusive rights holders. Opportunistic sharing arrangements in licensed bands could accommodate transient spectrum demand. (I know that this frequency is licensed to you, but you're not using it now, so could I use it for a few minutes while I'm passing through the area?)

3.3 Licensing

In licensed bands, the FCC specifies which entity is authorized to use which frequencies (or, more generally, electrospacetime volumes). The licensees use the frequencies to provide the specified services according to the rules that apply to the band allocation. In most cases, the license authorizes a specified signal at a given location or within a given area to provide the licensed service. In some cases (e.g., point-to-point microwave links), the associated receiver site is also licensed (which gives it legal protection from interference).

This licensing approach requires the FCC to determine which entity should get the license, and to decide whether a license is actually available to be awarded. We will dispense with the issue of “who” first, mainly because details of who gets a license are not particularly important for a paper on the technical issues of spectrum management, and then concentrate on the potential for a new license.

3.3.1 License Allotments and Auctions

The history of FCC licensing approaches illustrates the pressures towards market allocation of spectrum rights as demand for spectrum access shifts out.¹⁰ The original licensing approach allowed all eligible applicants to obtain ~~one~~ one license, usually for a small registration fee. The band allocation contains specific eligibility criteria for some licenses and a description of the procedures to obtain a license. In earlier times, frequencies were often “allotted” to specific types of users. Thus, in FCC LMR bands, the FCC allotted different groups of frequencies specifically to the lumber industry, the movie industry, local transit services, public safety, local government, railroads, and oil companies. This led to inefficient allocations of resources because in a given geographical area all of the licenses in one allotment might be fully devolved, while frequencies in other allotment categories remained unused. More recently, most of these narrow categories have been combined into more general categories. Nevertheless, some bands retain special categories of eligible licensees.

Comment [RJM1]: “shifts out”? ...shifts outside of existing allocations.

As frequencies became more heavily used and licenses became more valuable, the excess demand prompted the FCC to develop a means to select which eligible applicants receive a license. The FCC set up comparative hearings, an expensive and time-consuming process, to determine which one of several competing applicants would provide the best benefits to the country at large. The comparative hearings were sometimes called “beauty contests” or “liar’s contests,” depending on one’s point of view.

When comparative hearings finally became unworkably complex, in the 1980s the FCC began license lotteries. Predictably, individual applicants would hire “application mills” to fill out hundreds of license applications in hopes of being awarded a valuable license, and the FCC would receive tens of thousands of applications for a distribution of licenses. Although this process did work in the sense of expeditiously awarding licenses, it had very few other redeeming virtues. In particular, it raised the issue of whether the lottery winners could immediately sell their licenses to others (not actually permitted by law) who could actually use the licenses productively.

The last major FCC change in granting licenses was to auction licenses, beginning in the mid 1990s. Auctions have several major advantages, including raising substantial federal revenue, transparency, and an efficient initial allocation to companies that would put spectrum into the highest-valued uses. In general, the auction process has worked well (not perfectly, but much better than the earlier alternatives).

Auctions, however, have not addressed inefficiencies in the structure and substance of that which is auctioned. First, the FCC has not applied auctions to all spectrum blocks. Second, profound transactions costs impede the secondary markets for spectrum licenses, despite the good intentions of the FCC’s Secondary Market Initiative.

3.3.2 Spectrum Scarcity and Electrospace Volumes

¹⁰ For more on the history of license allocation policy, see Jonathan E. Neuchterlein and Philip J. Wieser, Digital Crossroads, American Telecommunications Policy in the Internet Age, MIT Press, Chapter 7.

In some frequency bands and applications, little spectrum scarcity exists. To obtain a license, one may need only apply for it. For example, in some of the fixed services (point-to-point microwave bands), the typical way to obtain a license is to hire a competent consultant (such as Comsearch) to design the propagation path, select a frequency, check the existing licenses for possible interference, and publish the tentative new license details. If no one objects within a certain number of days, the license goes into effect. Likewise, in the 70-, 80-, and 90-GHz bands, licenses are essentially free for the taking and only require registration of the desired frequencies. However, virtually no commercial equipment or components using such high frequencies yet exist, so the only demands are speculative. These frequencies are likely to be used commercially only for short-range, high-bandwidth, directional links. Thus when they are eventually in demand, the useful electrospacetime volumes are likely to be extensively partitioned geographically.

In bands that are already heavily in demand, the ways in which regulators initially partitioned the electrospacetime vary greatly. In some bands, such as the original VHF TV band, the FCC determined the total number of licenses in each city at the time that the entire band was allocated. These initial channel assignments ensured that co-channel licenses had sufficient separation to prevent any harmful interference between them given the technology at the time. The FCC conducted studies on adjacent-channel assignments, assuming that consumer TV receivers would have certain adjacent-band rejection characteristics. In the UHF TV bands, channel protection studies involved as many as three or four channels on either side of each licensed frequency. The FCC determined every possible transmitter license, and applicants would apply for the license for a specific transmitter, easily identifying available licenses by noting which ones had not yet been issued.

In most frequency bands, however, the total number of licenses and their locations is not predetermined. Instead, the FCC uses algorithms to determine whether a requested license can be granted, based mainly on whether it would interfere with the existing licenses that have already been granted. Based on engineering studies of the specific service and frequency band, the rules might say, for example, that the requested land mobile radio station must be at least 60 miles from all existing licensed stations using the same center frequency and at least 25 miles from any existing station using adjacent-frequency channels. The FCC may adjust these minimum separation distances up or down, depending on the details of the proposed radio site, antenna height and gain, and transmitter power. For point-to-point microwave systems, the FCC would also include antenna beamwidth, pointing angle (azimuth), and detailed topographic data. Using these rules, information about incumbents, and the information in the license application, the FCC determines whether or not to grant the new license.

Importantly, in this approach the partition of the electrospacetime into licensed volumes is a function of the chronological order and exact location for which new licenses are requested. For example, assuming a 60-mile minimum separation distance rule, the FCC could grant a new license for a station midway between two stations separated by 180 miles. Granting this single license would block granting any additional licenses between the original two stations. However, if the requested new station had been only 60 miles from one of the original stations, the remaining 120-mile gap would have permitted the licensing of an additional station in the middle of the

gap. Therefore, the order in which the FCC receives license requests affects the number of stations that fit into the 180-mile space between the existing stations. This procedure makes it difficult to know in advance how many licenses the FCC can grant in a band. It is also difficult to know the degree to which this procedure grants an optimal number and location of licenses.

As market demand and technology changes, the FCC may revise its algorithms for granting additional licenses. For example, in some mobile radio bands, the high demand for additional licenses and the availability of audio compression technology drove the FCC to change the bandwidth of channels in the band, from 30 kHz channels to 15 kHz channels to additional 15 kHz channels offset by 7.5 kHz to fit between the original 15 kHz channelization. In addition, the FCC may decrease the minimum required separation distances on the basis that communications with a little interference is better than no communications at all, or possibly assign a given frequency to multiple sets of users. Improved receivers with better adjacent-channel rejection and more-efficient digital modulations have also allowed regulators to decrease the bandwidth of channels and to squeeze the narrower channels closer together. In addition, improved propagation modeling allows planners to move from very approximate rule-of-thumb algorithms (e.g., 60-mile minimum separation distance) to more exact parameters based on topographical data. Thus, the combination of increased market pressure and improvements in technology has (in some bands) prompted the FCC to greatly increase the number of grantable licenses.

3.4 Interference

A key part of the design of a spectrum-dependent application is ensuring that the application functions properly (transmitting and receiving signals as required), and, in doing so, does not violate the rights of others to ensure the same satisfactory performance for their own systems. Indeed, the threat of interference is generally recognized as the major reason why spectrum users need distinct rights of access, even in unlicensed environments.

Interference is *the degradation of the performance of a receiver by the presence of an extraneous radio signal*. Interference occurs only within a receiver, although the extraneous signal that causes the interference is sometimes (but inaccurately) called “interference” or “interfering signal.” Interference -- by definition -- requires a receiver and a desired signal whose reception could be degraded. Importantly, this definition does not establish responsibility for the presence of interference or the mitigation thereof. There are two distinct types of interference, co-channel and out-of-band. Co-channel interference is caused by unwanted signals present at the same frequency as the desired signal. Out-of-band interference is caused by (usually strong) signals at frequencies that are different from the desired signal.

3.4.1 Co-channel interference

Co-channel interference is caused by an unwanted signal at the same frequency as the desired signal. This includes a signal that is centered on the tuned receiver frequency, as well as sidebands or other spurious radiated components belonging to signals whose main energy is at other frequencies, but which have components at the tuned frequency of the receiver. Since the

receiver is most sensitive to signals at its tuned frequency, relatively small unwanted signals can cause co-channel interference. Typically, interference will occur whenever the ratio between the received power in the desired signal, S , and the received power in the unwanted (interfering) signal, I , is less than K , where $K(\text{dB}) = 10 \log (S/I)$. Depending on the modulation, K is typically 4 to 20 dB, meaning that co-channel interference might occur whenever the desired signal is larger than the unwanted signal by less than a factor of 4 to 20 dB.

Co-channel interference often occurs near the edge of the coverage area of the radio system, where the desired signal, S , is relatively weaker, and the unwanted co-channel signal, I , coming from transmitters in adjoining areas are relatively stronger. Co-channel interference can also be caused when two mobile users on the same frequency try to transmit at the same time. In general, co-channel interference is relatively well understood, and the FCC licensing process prevents licenses for the same frequency from being assigned “too close” together geographically. The FCC generally assumes that receivers will have imperfect rejection of signals transmitting on adjacent channels (because of unwanted transmitter sidebands, as well as imperfect receiver bandpass filters). Therefore, it adopts similar rules regarding how close adjacent channels can be licensed. However, as bands become more crowded, the FCC may decrease the spacing between users licensed to use the same (or adjacent) frequency to accommodate more users, causing some additional co-channel interference in the absence of improved receivers.

3.4.2 Out-of-band interference

Out-of-band interference occurs when one or more strong signals at a frequency different from the tuned frequency of the receiver cause some of the electronic components in the receiver to partially overload, impairing its performance. In particular, intermodulation (IM) products derive from combinations of strong signals that enter the first RF amplifier stages in the receiver. If these signals are sufficiently strong, they will cause slight overloading and distortions in the RF amplifier. These distortions generate a large number of spurious signals at additional frequencies. If any of these additional frequencies happen to coincide with the frequency to which the receiver is tuned, IM interference can occur. IM interference is harder to predict than co-channel interference because it is caused by combinations of signals not directly associated with the desired signal. Moreover, by assumption, these other signals are operating within the conditions in their licenses. Therefore, there is usually less remedial action that can be taken against the signals causing IM interference. Fortunately, IM interference typically occurs only in very strong signal environments which are usually confined near to transmitter sites.

Engineers apply two main approaches to reduce IM interference. The first is to construct receivers that are more resistant to IM interference by designing RF amplifiers that can handle more total signal power without distortion. However, such amplifiers use more electrical power, and this can substantially compromise the utility of battery-operated portable equipment. Another approach is to reduce the probability that strong signals will get to the first RF amplifier by incorporating a narrower bandwidth RF filter to protect the RF amplifier. A narrower RF bandwidth proportionately reduces the chance of strong signals within the bandwidth. However, narrower RF filters reduce the range of frequencies over which one can tune the receiver,

possibly decreasing the usefulness of the receiver while increasing its size or manufacturing costs.

The other approach to mitigate IM interference is for spectrum planners to carefully plan the radio environment so that unwanted radio signals are less likely to cause out-of-band interference. Long-term spectrum planning can separate the applications involving high transmitter power from applications requiring large numbers of battery-operated receivers. However, historic band allocations are generally quite difficult to change. Moreover, it may be impractical for band allocation rules to keep up with the rapid changes in future technology that might improve or degrade out-of-band interference. In many IM interference situations, the problems are caused by specific combinations of local (i.e., co-sited) transmitter and receiver frequencies. In these cases, regulators or licensees can examine the frequency combinations and adopt a small change in transmitter or receiver frequencies that will eliminate the IM issues.

3.4.3 Interference protection sets the boundaries of spectrum rights

A sufficiently good receiver can eliminate almost all types of interference. For example, a directional antenna can often ~~reject~~ separate an unwanted signal at the same frequency as the desired signal interference. Likewise, a sufficiently poor receiver can experience interference in almost any environment.

In Section 3, we noted that an ideal receiver can successfully separate any two signals occupying distinct electrospacetime volumes. In theory, two parties could negotiate the boundaries of their respective volumes such that a cost effective level of receiver performance obtains. This classic Coasian bargain would be the natural result of internalizing the externality of unwanted emissions by establishing clear property rights. However, the FCC's band allocation rules (either explicitly or via other assumptions) assume an expected level of receiver performance and thus dictate the electrospacetime boundaries, including the separation distance between two systems using the same frequency, the bandwidth of the signal relative to the frequency spacing, and the transmitter power.

Under such an approach, as long as a receiver maintains a minimum level of interference rejection, any interference that actually degrades system operation is a violation by those who are emitting the interfering signal. On the other hand, if the receiver does not meet the required interference-rejection specifications, the interference impugns the receiver. Technically, "receiver-caused" interference and "interfering signal" interference may be identical, but the actual difference lies in the party whose rights prevail. For example, in a mobile radio system, co-channel interference is generally blamed on the "interfering signal" instead of the receiver. However, this assignment of blame assumes that receivers in a mobile radio band do not need directional antennas to separate out multiple signals on the same frequency. A future mobile radio system might possibly be expected to incorporate adaptive "interference-nulling" antennas; under these conditions the electrospacetime rights could configure differently. In contrast to the LMR scenario, a modern fixed microwave receiver (and associated receiving antenna) must be able to reject same-frequency signals on the basis of direction of propagation. Because fixed

microwave systems operate by much different rules than mobile systems, the corresponding expected receiver performance is very different.

3.4.4 Interference disputes and tenant's rights

In almost any interference situation, the first problem is to understand the technical cause of the interference. Close behind, however, is the problem of assigning responsibility for fixing the problem. Such assignments can be very difficult in contexts with no clear (i.e., specific, quantitative) understandings concerning the required level of receiver performance.

The FCC designs its rules regarding emission masks, transmitter separation criteria, and other licensing rules to prevent harmful interference. If all parties follow license rules, systems will usually operate without interference. But what if all parties are in compliance with all of the applicable rules and specifications, and interference still arises? Under these circumstances, the principle of *tenant's rights* applies.¹¹ This principle states that earlier-licensed stations at a site have precedence and that the most recent station must take the necessary action to eliminate whatever interference occurs as the result of its new presence.

This approach to establishing the boundaries of spectrum rights is enacted for federal government spectrum in NTIA rules. Section 2.3.7 of the NTIA Manual states:

“2.3.7 Spurious Emissions In principle, spurious emissions from stations of one radio service shall not cause harmful interference to stations of the same or another radio service within the recognized service areas of the latter stations, whether operated in the same or different frequency bands.

Providing appropriate spectrum standards in Chapter 5 are met, an existing station is recognized as having priority over a new or modified station. Nevertheless, engineering solutions to mitigate interference may require cooperation of all parties involved in the application of reasonable and practical measures to avoid causing or being susceptible to harmful interference.”

Section 5.0.1 of the NTIA manual states:

“5.0.1 Consequences of Nonconformance with the provisions of this Chapter. In any instance of harmful interference caused by nonconformance with the provisions of this chapter, the responsibility for eliminating the harmful interference normally shall rest with the agency operating in nonconformance.”

These two sections suggest a procedure for fixing the interference. One would first examine the emission spectrum of the interfering transmitter and the operating characteristics of the victim receiver to ensure that both meet the applicable standards. If either system is nonconforming, it

¹¹ The term “tenant’s rights” is also sometimes called “first-in-time is first-in-rights.” We have used the “tenant’s rights” terminology to emphasize the spatial and other aspects of the problem in addition to the temporal aspect.

must be brought into conformance. If the interference persists, the new or modified station bears the responsibility for any modifications needed to eliminate the interference.

The last part of Section 2.3.7 of the NTIA Manual suggests that, at least within the community of Federal users, a somewhat flexible approach to fixing interference problems may be useful. The flexible approach is useful because, although the most recent user may have the responsibility to fix the problem, the solution may be most easily accomplished by modifying one of the earlier systems. Therefore, less expensive fixes may be available by considering a broader class of solutions.

Several sections in 47 CFR describe the FCC approach to the problem. The rules for several services carefully define – typically in paragraphs (a) and (b) – an applicable transmitter emission mask, followed by paragraph (c) that says, for example,

“Should harmful interference be caused to the reception of other ... stations by out-of-band emissions, the licensee may be directed to achieve a greater degree of attenuation than specified in paragraphs (a) and (b) of this section.”¹²

However, the regulation gives no indication of how much additional attenuation the FCC may require or what applicable principles or processes the FCC should use to decide. Other rules in certain services extend the authority of the FCC to require adjustments in effective radiated power, antenna construction, and antenna polarization.¹³

Neither the NTIA Manual nor 47 CFR spells out tenants rights in detail because detailed band allocation rules historically prevented most interference. However, new approaches to allowing licensees to negotiate levels and kinds of interference could justify more clarity in interference liabilities and the procedures to resolve conflicts.

3.5 Receiver Performance Standards

If better receivers can reduce interference, why not require higher-performance receivers? Such rules would surely decrease interference, allow more users per band, and generally improve spectrum technical efficiency and capacity. This logic is consistent with a command and control regulatory approach because without price signals, market forces, and flexible rights, an efficient level of receiver performance is difficult to ensure. However, mandating receiver technologies is generally economically inefficient, and we argue here that they are unnecessary and possibly counterproductive even in a command and control context.

¹²47 CFR §73.44 (c), for example. This regulation applies to AM broadcast radio. Similar regulations can be found in §73.687 (television broadcasting, particularly channels 14 and 69 which are adjacent to mobile radio bands).

¹³47 CFR §21.106, §21.107, §21.109, and §21.110

Since the engineering of an allocated band (including channel spacing and frequency re-use distances) assumes a set of receiver performance factors, licensed systems can normally operate as planned if the receiver performs at least as well as the design criteria for the band. In addition, as discussed in the section above, one criterion for deciding who is responsible for fixing interference problems is whether the victim receiver meets the expected performance specifications. Further, 47 CFR suggests that the victim receiver must be at least “average” in its performance to qualify for protection against interference. Therefore, users may be able to provide operationally satisfactory service with a cheap receiver, but choose more expensive receivers to be eligible for the protection from interference from illegal signals.

Another reason a specific performance level is undesirable is that optimal receiver performance can vary considerably across different users in the same band, depending on factors like the distance between the receiver and nearby transmitters and the strength of signals on adjacent channels. In most cases, receiver owners are aware of their emissions environments, including the presence of strong signals and the priority given to interference-free operation of the system. Moreover, if a receiver owner deploys an inadequate receiver, he or she bears the full cost of that problem. Unlike non-conforming transmitters (which can cause interference to other users), poorly performing receivers create no external costs to others. Further, a receiver owner may rationally choose to deploy a low quality, low cost receiver if the risk to service quality is worth the savings in receiver costs. At best, regulating receivers in this context inefficiently protects users from themselves. At worst, receiver performance standards may make matters worse; regulators have less information about users’ specific circumstances than the users themselves, and in any case users’ receiver performance needs are heterogeneous.

Another reason to avoid receiver performance standards is that ordinary technical standards are too simplistic for today’s technologies. In bands in which receiver performance rules would be most likely, the standard approach would be to set fairly simple hardware specifications like bandwidth and adjacent channel rejection (possibly including a selectivity mask), noise figure or sensitivity, intermodulation specifications, and maximum local oscillator radiation. However, modern systems can deploy more advanced interference rejection technologies. For example, a modern radio system may be smart enough to choose another frequency where the signal environment is less challenging or possibly to use complex signal processing to avoid the interference. One would not want to preclude these creative and sophisticated approaches with traditional regulation, but writing more general rules would be tricky.

Our objection to performance standards applies strictly in the policy domain. Where receivers are part of a system where the faulty operation of one receiver can cripple the performance of other parts of the system, the other parts of the system have a legitimate interest in the performance of every receiver. Public safety, defense, and many other large radio networks are examples of such systems. However, in these cases, defining the performance of receivers (along with all other system components) is properly the role of the system designers -- not the role of spectrum regulators.

3.6 Duration of rights

The electrospacetime model we presented in Section 2 is fully explicit in the time period over which a user controls the rights to access spectrum; indeed, time is one of the key dimensions of the electrospacetime. The system of FCC licenses for non-federal entities and NTIA assignments to federal agencies stands in contrast to this clarity. Most FCC licenses have a specified multi-year duration, but carry the strong supposition (but not a guarantee) that the FCC will renew the license when it expires. This renewal process could repeat indefinitely, approximating permanent rights but introducing important risks that regulators could encumber the renewal process with reviews and extra requirements.

NTIA regulations require it to review spectrum assignments to federal agencies every five years, which it does not typically do. Indeed, NTIA has delegated to some agencies the day-to-day responsibility to regulate their use of certain spectrum bands, generally where the agency is the largest user in the band and where technical details of agency systems are quite specific to the use of that band. For example, the Federal Aviation Administration (FAA) has the responsibility to manage the 2700-2900 MHz band, which contains S-band airport surveillance radars operated mostly by the FAA.

In the past, the lack of well-defined temporal rights has served mainly to ensure that spectrum that is no longer needed for its original purpose will remain unused, instead of being promptly re-used for other purposes. Until recently, FCC licensees had no rights to reassign that spectrum to other users. Therefore, the only choice they had was between keeping possession of the unused spectrum rights or returning the license to the FCC. Recent FCC policy changes have encouraged licensees to rent or lease unused licenses to other users.¹⁴ However, this policy change tends to undermine past philosophical fictions that licenses conveyed no property rights, but only conveyed the right to a specific user to use a frequency for a while for a single, very specific purpose.

In general, FCC licensees and federal agencies seem to live with ambiguous temporal rights reasonably well. We argue, however, that clarity in the temporal dimension of spectrum rights is likely to become increasingly important in the future, particularly in the context of dynamic spectrum access, market driven resource allocation, and spectrum leasing and other secondary market activity between and among federal and non-federal entities.

3.7 Command and control conclusions

The command and control approach has many advantages. It allocates frequency bands with optimal rules for diverse specific services, such as LMR, broadcasting, radars, and point-to-point microwave. Most of these allocated bands remain well designed from a technical perspective, allowing reliable services for many users with a minimum of interference. Further, detailed allocation rules provide a template for radio system designs, and users can be assured a good degree of performance since incumbents are well-protected from new entrants.

¹⁴ FCC Report and Order and Further Notice of Proposed Rulemaking (FCC 03-113). May 15, 2003

However, the command and control approach incurs significant, but uncertain opportunity costs. In no way does it guarantee the efficient allocation of spectrum resources across and within allocated bands. It substitutes the judgments of regulators for market-driven outcomes, and explicitly prohibits Pareto-improving transactions and resource reallocations. Despite the FCC's Secondary Market Report and Order, burdensome rules still apply to many license transactions, and among the things *not* included in most licenses are rights to change the use of the license in any way, including changing transmitter location, receiver location (if there is a specific receiver location), bandwidth, modulation, and services provided. These inefficiencies likely incur profound dynamic costs by impairing technological change by increasing the risk that new technologies cannot find sufficient accommodation in the established band plan.

Although regulators eventually match most proposed spectrum uses with a frequency band in which they can operate, it can take years and require millions in legal and consulting fees. In some cases, regulators can slightly modify existing band allocation rules to accommodate the new technology, for example by changing the bandwidth/channelization of a mobile channel from 25 kHz to 12.5 kHz. In other cases, regulators could relocate incumbent licensees in a relatively unpopulated band and redevelop the band with new allocation rules. Such a process converted the old 1850-1990 MHz point-to-point microwave bands to the new Personal Communications Systems (PCS) bands. However, current approaches to accommodating new services are cumbersome, costly, and result in great delay and forgone consumer and producer surplus.

Arguably, NTIA's approach must remain more command and control than the FCC's. NTIA has some frequency bands with numerous unique systems designed for military or continuity-of-government missions, so NTIA must engineer each one into a band independently. In addition, certain frequencies must be available in the event of a national emergency with a certainty exceeding any commercial service quality requirements. However, new policy and technological approaches could possibly allow productive use for some of the "reserved" spectrum during the long periods between emergencies.

Fortunately, regulators with command and control authority can use their discretion to implement band rules quite different from traditional exclusive licensed bands. In the subsequent sections we explore these possibilities.

4. ALTERNATIVES TO COMMAND AND CONTROL EXCLUSIVE RIGHTS

Figure 6 below places alternative spectrum management techniques along two axes. The horizontal axis represents the degree of flexibility users have in designing their radio systems and

providing services, going from the least flexibility on the left to complete flexibility on the right. The vertical axis represents the degree to which those who access spectrum have exclusive rights to that spectrum as opposed to default sharing with others. Users near the top of the figure have exclusive rights to their spectrum, while users near the bottom obtain their spectrum from a common pool used by others.

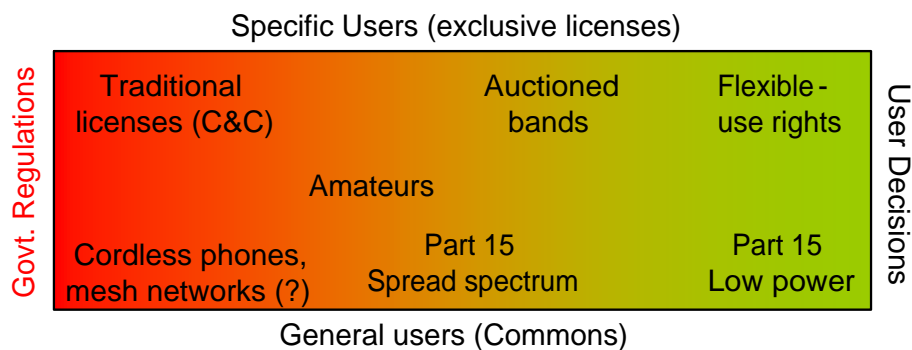


Figure 6. Spectrum management techniques along two continua

Traditional exclusive licenses lie in the upper left, with detailed federal rules but high degree of interference protection. The lower right-hand corner includes the unlicensed low-power spectrum allocation (for example as described in 47 CFR Part 15), which allows users to operate low power devices without a license.¹⁵ Proliferating unlicensed devices currently offer countless short-range wireless services, including garage door openers, wireless routers, cordless phones, and remote control toys. In the lower left corner of Figure 6 lies a class of systems that, although not licensed to specific users, follows very specific federal regulations. For example, Family Radio Service (FRS) is a service intended to provide families and other non-commercial users an unlicensed medium-range mobile radio service on designated frequencies/channels. Regulations tightly specify the parameters of the service. Spread spectrum systems also operate in unlicensed bands. More specific regulations govern their use than low-power systems, including specifications on frequency-hopping rates and number of frequency steps. The upper right corner includes systems using flexible-use spectrum rights. These systems use frequencies that are exclusively licensed to specific users, but licensees have broad discretion about how they employ and share their frequencies.

We will return to thinking about spectrum rights approaches along these two axes as we go through three major alternatives to exclusive licenses with detailed service rules: unlicensed use (Section 5.1), opportunistic use (Section 5.2), and flexible spectrum rights (Section 5.3).

4.1 Unlicensed spectrum

¹⁵ We use the common term “unlicensed” to refer to spectrum in which FCC does not require users to have a license.

Low power, highly localized services are not well-suited to traditional exclusive licenses. Each device occupies such small electrospacetime volumes that the cost of individual licensing would swamp the value of the electrospacetime used. In general the FCC sets two major conditions for devices that operate in unlicensed bands. First, the FCC approves the equipment for use in the band, and it generally limits transmitters to very low power. Second, users have no presumed protection from interference. However, given the low power limits, interference is usually caused only by other users operating in the immediate vicinity. Therefore, interference is not usually an issue.

The small bubbles of electrospacetime occupied by unlicensed devices pose a minimal risk of interference to licensed systems, particularly licensed systems used by anyone other than the user of the unlicensed device. One can think of unlicensed spectrum as the regulator granting bubbles of individual electrospacetime volumes to unspecified nameless users of authorized devices, bubbles which together form a virtual foam of user rights. The power limit of the devices dictates the radii of the individual bubbles. Thus an unlicensed low-power band is not truly a commons in which “anyone can do anything,” but rather an *a priori* partitioning of the electrospacetime into small parts that are unlikely to intersect, at least until devices proliferate to the point that congestion arises.

As electronics technology has improved, the FCC has allowed some higher power devices in the unlicensed electrospacetime, provided that it includes other means to reduce interference. For example, some Part 15 devices include “listen-before-talk” hardware and software that searches a frequency range for existing signals and moves to operational frequencies that are not already in use. As technology matures and congestion increases, more elaborate interference-avoidance capabilities in unlicensed spectrum may become cost effective, including “listen-before-talk” protocols, dynamic frequency selection (DFS), and other techniques.

The unlicensed space is most certainly another kind of command and control approach. It mitigates interference through low power or other technical constraints. A number of inefficiencies can arise from this scenario. First, unlicensed devices mitigate the risk of interference even when there are no other users around. For example, the range of a cordless phone is the same (bounded by regulation) no matter whether one operates it in the center of a 100 acre farm or in a studio apartment in the city. This lack of context-specific protocols leaves spectrum inefficiently underutilized in some instances. Second, different technologies may be better suited to different sets of unlicensed rules, and there is no way for the FCC to ensure that its portfolio of different rules in different unlicensed bands is efficiently designed.

For example, unlicensed advocates have proposed various types of self-organizing wireless mesh networks as a possible new class of radio systems that could automatically recognize and connect to one another in adjacent areas. This self-organized network would allow users to pass packet data from system to system until the intended destination was reached. Thus, a mesh network could operate over long ranges while having the frequency reuse efficiency of short-range systems and adding extra capacity at the same rate that more users are added. Although the large-scale feasibility of such systems remains to be demonstrated, they show promising

potential. It might result that wireless mesh networks perform much more efficiently if specific rules/protocols are enforced to prevent certain types of interference from other unlicensed uses. However, the FCC faces the conundrum of allocating spectrum to unlicensed uses with rules for technologies that are not yet mature, so one cannot know the optimal rules for their deployment. Yet without spectrum in which to operate, investors will not deploy the technology.

Unlicensed spectrum raises another policy question. If unlicensed devices are allowed to operate without charge in a non-licensed band, some argue that this represents a public subsidy compared to other radio services where users are required to pay for spectrum licenses. The degree to which it is economically efficient for the government to charge for spectrum access depends, among other things, on the transactions costs of collecting revenue and the net social gains that derive from the revenue.¹⁶

4.2 Opportunistic Access

Almost all spectrum experiences periods of relative idleness as the supply and demand for radio communications fluctuates over time. In addition, most licensed uses of spectrum have unused “white spaces” around the edges of their service area caused by signal propagation across irregular terrain and power limits to prevent interference to their neighbors. Likewise, regulations continue to reserve many older frequency bands for their allocated purposes, even when these bands are only partially filled by licensed users. Even in frequency bands that are filled in crowded urban areas, large amounts of “white space” may persist in less crowded rural areas.

In this context, some may argue that dynamic-use systems could legitimately operate as opportunistic overlays, able to access any spectrum used by any system in any management regime, so long as they produce no harmful interference to the licensed incumbents.¹⁷ In theory, one could argue that this is consistent with current policy. Unlicensed spectrum use carries no guarantee of interference protection, and current exclusive licenses guarantee a level of freedom from interference, not full control over the electrospace volume.

At least two concerns arise with this logic. One centers around the problem of how to provide certainty that no interference or other costs/inconveniences will be caused to licensed users, not just now, but in the future as licensed users modify their spectrum use within their licensed rights. The second is whether an opportunistic user should be able to freely access the “idle” spectrum without the consent of or compensation to exiting licensees or those to whom licensees may have leased out spectrum access rights. Naturally, licensees may be uncomfortable with unauthorized “trespass” on their licensed rights, even if they cannot prove they have experienced

¹⁶ For a detailed discussion of the economics of raising government revenue from spectrum access, see “Spectrum Auctions: Distortionary Input Tax or Efficient Revenue Instrument?” by Adele C. Morris, *Telecommunications Policy* 29 (2005), pp. 687-709.

¹⁷ An “overlay” refers to opportunistic spectrum use by a system using relatively high transmitter power. Overlay devices must employ suitable techniques to determine that they will not cause unacceptable interference to existing systems before they can transmit.

harmful interference. This concern highlights the importance of clarity in the degree to which licensees can exclude all signals of others or merely excluding signals that would cause interference to the licensed systems. Section 6 below argues that a regulatory system that employs an electrospatial approach to expressing spectrum rights would allow consenting licensees to permit opportunistic use of their spectrum with the minimum of transactions costs.

Use of relatively high transmitter power by opportunistic overlay transmitters greatly increases the risk of causing interference, in part because the potential interference range is greater and in part because it is more difficult to infer the presence and the technical operating parameters of the more distant potential victim receivers. A number of researchers have proposed a wide range of methods to obtain better data on the presence of distant victim receivers. These include such techniques as:

Formatted: Highlight

1. Include more sensitive search receivers in opportunistic devices.
2. Network opportunistic devices to get search information from distant devices.
3. Incorporate database information on licensed devices in opportunistic systems.
4. Require licensed receivers desiring protection to transmit status to real-time data base.
5. Include access to propagation modeling with GIS, license data base, real-time status.
6. Introduce a band manager with access to (5) who would grant or deny permission to use.
7. Victim receivers could measure test signals and report on test signal from opportunistic user.
8. Interference temperature calculations at victim receiver based on everything above.

None of these techniques is foolproof or easy. Some require substantial cooperation (and expense and inconvenience) on the part of licensed devices, decreasing the value of a license. A band manager would presumably need to be paid for its services, which payment might (or might not) include payments to licensees whose frequencies were “borrowed.” Presumably, smaller parts of the list might be used for low-to-medium opportunistic transmitter power, while more sensing/testing would be required for higher transmitter power. Similarly, penalties for causing interference might escalate as opportunistic users increased transmitter power, knowingly increasing the risk to licensed users. On the other hand, some licensed users might choose to invite opportunistic sharing (for a fee), while other licensed users might prohibit opportunistic use on their frequencies.

Present trends in the uses of wireless applications suggest that many more future wireless systems will provide services to users who are constantly changing location, services desired, services offered by others, and systems available to support these services (sometimes pedestrian, or moving in and out of offices, cars, etc.) Thus, the traffic to each user and the systems available to support it may be changing rapidly. Under these conditions, it is less likely that any single license owned by the user for any specific electrospatial region will efficiently match the communications requirement. Either the licensed large electrospatial region will remain unused most of the time (thus, a source of white space for other users), or the user will be almost continually outside the boundaries of a smaller electrospatial region (thus, becoming a potential consumer of white space belonging to others). Clearly, some mechanism would be useful to help

match the static/localized qualities of an electrospace license to the more dynamic/dispersed traffic demands of the user. Opportunistic use is one way to make productive use of these temporarily/geographically mismatched white spaces.

5. AN ELECTROSPACE APPROACH TO FLEXIBLE SPECTRUM RIGHTS

This section explores how regulators could use an electrospace approach to delineate spectrum access rights that are flexible, exhaustive, and economically efficient. We also address practical problems that arise. We draw much of the material in this section from earlier papers by Matheson.^{18,19} It is also very consistent with Kwerel and Williams[2006].²⁰

In our suggested regulatory regime, licensees would acquire and use rights to access spectrum that are expressed as disjoint electrospace volumes. In this system of “ideal spectrum rights,” interference and the allocation of the resource generally are governed by the following fundamental rule:

1. Licensees must keep all signals within their respective licensed electrospace volumes (including frequency band, geographical area, or authorized time of operation). In particular, all signals must have a power level of less than E_0 outside their electrospace region.

The rule simply reiterates the boundaries of the electrospace volume as the rights of spectrum access. Section 6.2 addresses the practical setting of E_0 .

This idealized approach is characterized more fully by the rules it does *not* include. Signals within the electrospace volume have no limitations on power or field intensity. No regulatory constraints on receivers or other technologies apply. No regulatory constraints on the services offered with the spectrum rights apply. No specific protection from interference (other than the electrospace parameters on other licensees) applies. Licensees may aggregate, divide, lease, or transfer their rights via secondary markets and private as they see fit without prior approval by regulators, and they may partition their rights along all electrospace dimensions. Licensees will presumably design their radio systems so that they can operate in the presence of small, unwanted foreign signals (less than or equal to E_0).

18. R. J. Matheson, “Flexible-use spectrum rights - a tutorial,” Proceedings of the International Symposium on Advanced Radio Technologies, March 2003. NTIA Special Publication SP-03-401.

19. R. J. Matheson, “Principles of flexible-use spectrum rights,” Journal of communications and networks, Volume 8, Number 2, June 2006, pp. 144-150. (ISSN 1229-2370)

²⁰ *Defining Spectrum Rights*, a presentation at the NTIA Workshop on Improving Spectrum Management Through Economic and Other Incentives, Feb. 28, 2006.

In this approach, the role of the government is to devolve initial allocations of electrospacetime volumes, to manage a public database of spectrum rights, to enforce those rights, and to mediate disputes when necessary.

We argue that this approach allows licensees to efficiently manage interference with the maximum of flexibility of spectrum use. Doing so would allow robust and liquid markets in electrospacetime volumes to emerge. The issues we discuss here concern only the delineation of spectrum rights and obligations. We do not address the means by which the user obtains those rights.

As long as all licensed users of non-overlapping electrospacetime regions follow the fundamental rule, each licensed user can operate without interference, no matter what the other compliant licensed users are doing within their own non-overlapping electrospacetime regions. Since we have made no assumptions about the actual sizes of the various users' electrospacetime regions, there is no reason to suppose that changing the size or number of electrospacetime regions would cause any additional interference. One must ensure that any partition of an electrospacetime volume creates truly disjoint sub-volumes, but other than that there is no reason to regulate secondary market transactions.

In the following sections we explore ways in which our "ideal rights" approach runs up against practical realities. Some of these practical realities mean we need to add addenda rules to our fundamental rule, striking careful balances across tricky tradeoffs.

5.1 Practical Electrospacetime Rights

The most serious limitation on the practical application of the electrospacetime model to flexible-use spectrum management is that the electrospacetime model assumes that all receivers can reject unwanted frequencies (i.e., signal power at frequencies outside of the nominal receiver bandpass), have infinite dynamic range (strong out-of-band signals will not cause intermodulation products or gain compression), and contain directional receiving antennas to reject signals coming from unwanted directions.

Thus the ideal spectrum rights world assumes receivers can separate any two signals that differ in at least one of their electrospacetime coordinates. This means that all interference is (for all practical purposes) always caused because the receiver is not good enough. Unfortunately, the required "good-enough" receiver for some circumstances might be extraordinarily complex and expensive, and it might not be achievable using today's technology.

Practical policy making must grapple with the tradeoff between the benefit to licensees of allowing them to access the full range of rights in their electrospacetime volume versus the possibly large costs to others of rejecting unwanted signals. Where parties can negotiate, the optimal approach is to establish clear property rights and allow the parties to construct a market-based resolution to the interference problem. Later in this section we discuss where we think this would work. In some cases – out-of-band interference for example – the number of affected

parties could be too large to allow efficient negotiations, so regulators may have to set rules that broadly strike a reasonable balance of interests, recognizing that updating is necessary in any areas of rapid technological change.

In the case of ideal receivers, users would only need to worry about foreign signals that illegally intrude at the frequency of the desired signal – so-called “in-band” interfering signals. These in-band intrusions would be limited by licensees observing the electrospatial rules that control “ E_0 .” Unfortunately, an important characteristic of practical (i.e., non-ideal) receivers is that they can experience interference even when no unwanted signal is actually present at the tuned receiver frequency. Strong signals at close-in frequencies or very strong signals at frequencies further away from the tuned frequency can also cause receiver distortions that are seen as interference; this is the “out-of-band” interference discussed in Section 4.4.2. This interference could be resolved by either ignoring the practical problems of receivers in limiting out-of-band interference and sticking to a pure electrospatial volume, or by limiting the power levels of licensees *even within their licensed electrospatial volumes*.

Most analysts suggest that the most cost effective approach is to supplement the ideal flexible-use rules with a limit on transmitter power. This produces a more benign signal environment that allows the use of less-expensive “good-enough” receivers. We are implicitly assuming that the costs of restricting licensees to lower transmitter power are lower than the benefits to other receiver owners of the more benign signal environment. In reality this is an empirical question that regulators must assess.

Likewise, the transmitter power limits in a given band and across bands should be chosen to maximize overall net benefits – balancing the benefits from less-expensive receivers with the disadvantages of more-restrictive limits on transmitter power. Regulators can choose different limits for different bands to provide a portfolio of options for different types of systems.

Thus, to our first fundamental rule, we add:

2. All signals must be no stronger than a power limit (E_{\max}).

We discuss several possible ways to apply the limitations on transmitter power in more detail in section 6.3 and discuss the way that these limitations scale when electrospatial regions are aggregated or divided in section 6.5.

5.2 Interference

A major advantage of the flexible-use rules (compared to dual-space spectrum-regulation models that include receivers, as described in section 3.6) is that the flexible-use rules lead to much less legal ambiguity about who is responsible for fixing interference situations. Assuming licensees obey the two flexible-use rules, receiver owners are completely responsible for solving their own interference problems. In an interference situation, the “victim” receiver owners have several ways to deal with the problem:

- a. Show that a specific transmitter is violating one of the applicable flexible-use rules (exceeding E_0 outside the licensed region or exceeding maximum power inside a region) and pursue compliance.
- b. Improve systems to eliminate the interference. Such changes might involve improving the victim receiver, increasing desired transmitter power, or adding better error correction.
- c. Tolerate the interference. This includes changing operating procedures, restricting the operation to areas where interference is not a problem, or adapting to lower quality of service.
- d. Negotiate with the interferer. This negotiation is voluntary for both parties, and could result in the redrawing of the boundaries of electrospatial volumes, changing transmitter power, financial transfers, or all of these.

The practical flexible-use rules help to produce a sufficiently benign signal environment such that licensees can deploy less costly receivers. Users have complete flexibility of exactly which receivers to deploy. They have much better knowledge of their requirements than regulators and are highly motivated to optimize cost effectiveness.

5.3 Setting E_{\max} and E_0 Limits

High signal levels within the relatively wideband first RF stages in a practical receiver are usually the major cause of out-of-band interference. Therefore, placing a limit on transmitter power (or EIRP) will reduce the occurrence of strong signals in the radio environment and make less-costly receivers feasible. Limiting transmitter power would not be expected to eliminate all out-of-band interference. However, out-of-band interference would tend to be limited to the less likely circumstances in which the victim receiver is located very close to a transmitter tuned to a nearby frequency. Therefore, the use of practical flexible-use rules will tend to allow interference-free operation in more locations using cheaper receivers.

Out-of-band interference to receivers results when a receiver is located in an area of strong field strength from an unwanted radio signal. This “high-field strength” condition results from a combination of three factors:²¹

1. High transmitter power,
2. Transmitter vertical antenna patterns that produce high-field-strength signals on the ground in locations that are open to the public, and
3. The presence of a susceptible receiver in the high-field-strength, open-to-the-public areas.

²¹ In this discussion, we are assuming that out-of-band interference that results from co-sited transmitters and receivers will be resolved by a site manager or by negotiations between the directly affected parties, without involving a large number of other parties.

Thus, out-of-band interference could be prevented by controlling any (or all) of the three factors. Instead of controlling the interference solely by limiting transmitter power, it would provide additional user flexibility to control interference by adjusting transmitter antenna patterns and/or by separating transmitter sites from likely concentrations of susceptible receivers.

Thus, the supplementary rule to protect receivers might include a limit on maximum signal field strength at ground level in public areas, instead of a limit on maximum transmitter power. This alternative rule would state that field strength at ground level must be less than E_{\max} , where E_{\max} corresponds to a maximum watts/m². This limit is not bandwidth-dependent, since the total power at the receiver input is usually what causes the problems, and the receiver front-end circuits will tend to be much wider bandwidth than most transmitters. This limitation must be met only in areas where there is a likelihood that susceptible receivers will normally be found. In some circumstances, it might be necessary to similarly protect additional above-ground-level outdoor locations where people are often found (e.g., elevated walkways, rooftop cafes on nearby buildings, etc.).

The maximum-field-strength rule would allow more flexibility in building a wide variety of radio systems, and it would protect receivers better than a limitation on transmitter power alone.²² Part of the cost of using higher transmitter power is that the field strength at ground level will need to be suppressed relatively more, so that it still meets the E_{\max} field strength limit. In an economic sense, this rule would tend to ensure that the higher cost of using a more powerful transmitter is borne entirely by the transmitter owner, instead of being partly externalized to large numbers of unrelated receiver owners.

The value of the parameter E_{\max} is completely determined by practical receiver technology; there is nothing theoretically absolute about these values. If a future change in receiver technology causes the performance of receivers to change substantially, this numerical value should also be expected to change. The recent development of various receiver-on-a-chip technologies have made receivers much smaller and cheaper, but not necessarily more resistant to strong unwanted signals. Future changes in receiver performance may result from much smarter receivers that figure out how to move to a better frequency or a better modulation, from receivers using digital RF or IF processing (where optimum bandpass filters can be synthesized), from room-temperature superconductors (producing very-narrow-band, very-high-Q, tunable RF filters that could reject many of the signals that otherwise would cause out-of-band interference in today's receivers), or from adaptive antenna technology (that could null out strong unwanted signals). On the other hand, software-defined radios (SDRs) and cognitive radios (CRs) may use receivers that are inferior in some ways to current receivers. The requirement to operate in many different frequency bands may curtail the use of passive RF bandpass filters, increasing the susceptibility to out-of-band interference.

²²Note that the interference to public safety LMR in the 800-MHz band in the US was caused partly by allowing apparently reasonable changes in antenna locations, without requiring changes in transmitter power. In this case, E_{\max} limits would have provided better protection from interference, while allowing more flexibility in use.

Rules for Setting and Scaling E_0

Although the electrospacetime model is critically based on a specified spectral power flux density limit, E_0 , which signals cannot exceed outside the licensed region, it is not obvious what numeric value to choose for E_0 . E_0 has the units of $W/m^2/MHz$, and it includes signals from a single transmitter coming from all directions (although most power will ordinarily come from the direction of the transmitter). Presumably E_0 will be chosen so that systems licensed in other regions will not receive interference from the signal. However, the minimum level of interfering signal for various types of systems varies over a wide range – perhaps 40-50 dB – depending especially on the gain of the receiving antenna and the required signal-to-interference ratio. Since all types of systems must be assumed to operate in a flexible-use band, which type of system should E_0 protect? One answer is that the selection of a specific value for E_0 might be done with the intention of making that band particularly suitable for various types of services; multiple bands could use different values of E_0 to efficiently accommodate different broad categories of services.

E_0 is scaled in terms of $W/MHz/m^2$. When geographical areas are added or subtracted from a region, the change merely affects the geographical position of the boundaries outside of which the signal must be suppressed below E_0 . Similar effects are applied to changes in the time and angle-of-arrival boundaries. When the frequency boundary is changed, the bandwidth of the signal leaking across geographical boundaries will change with the bandwidth of the primary signal, but the value of E_0 at any particular frequency will remain the same.

One complication of aggregating or combining electrospacetime regions comes from an understanding of what constitutes a “signal” that is entitled to leak E_0 into adjoining electrospacetime regions. The owner of a single electrospacetime region should not be entitled to leak more signal outside his electrospacetime region by simply claiming that a single electrospacetime region (and signal) had been divided into multiple electrospacetime regions (and signals) – each “signal” with a separate allowance for E_0 . The exact rules that define a single “signal” may be hard to state exactly, since independent signals can be combined in many interesting ways. If ten independent radio signals are combined and amplified by a single wideband power amplifier and transmitted from a single antenna, is the output one signal or ten signals? In summary, E_0 does not need to scale in any way under aggregation or division of electrospacetime regions. However, the arbitrary division of a signal into separate pieces to acquire a separate allowance for E_0 for each divided portion is not permitted.

Although a transmitter can radiate a large amount of power inside the licensed frequency range, the signal strength outside the licensed band must be less than E_0 . Presumably this condition must be met at all locations – even very close to a transmitting antenna, where the field strength is very high. At such locations, very high field strength inside the licensed frequency range would need to drop below E_0 immediately outside the licensed frequency range (possibly within a 12.5 kHz LMR channel) – requiring a very rapid decrease in signal strength over a small change in frequency. Therefore, the requirement that emissions outside a licensed electrospacetime region be less than E_0 , may need to be supplemented by an optional “relative-dB” emission mask

that provides a legal “safe harbor” in locations where the field strength is very high. This emission mask could be similar in required bandpass shape to existing transmitter emission masks. In locations where the signal strength is lower, the optional emission mask would usually require emissions to be lower than E_0 , thus limiting the use of the emission mask to a few locations near transmitters.

Rules for Scaling E_{max}

In terms of scaling maximum transmitted power along other electrospatial axes, the dimensions of time, space, or angle-of-arrival do not cause any difference in transmitter power scaling. Extending the geographical area of a region may allow more powerful transmitters to be employed, simply because the new geographical boundaries are further from the transmitter site, permitting more transmitter power without violating the leakage of signals above E_0 outside the new boundary.

Although the installation of additional transmitter sites within a region can increase the total power radiated at a given frequency, this will generally not increase the risk of interference to other users. The probability of interference from out-of-band signals is primarily related to the presence of *strong* unwanted signals, not by the total area over which a weaker unwanted signal is available. Therefore, there is no reason to limit the total power radiated by multiple sites, as long as the total power radiated by a single site is controlled.

Since the occurrence of out-of-band interference is mostly related to the total amount of signal within the very-wide-bandwidth electronic circuits at the receiver front end, it should be assumed that all of the energy from any transmitter at any nearby frequency will be available to cause out-of-band interference. Therefore, the E_{max} limit does not scale with transmitter bandwidth, but remains tied to a fixed maximum field strength. Presumably, the cumulative power from multiple transmitters at the same site should be included within this limit, with some rules for requiring compliance by any group of multiple transmitters that cumulatively violates the field strength limit.

5.4 The Problem of the Height Dimension

The electrospatial defines locations in a 3-dimensional way. This means that the spatial description of any electrospatial region is a 3-dimensional solid volume, including height (or altitude) as well as latitude and longitude. However, the height dimension is problematic in some ways. For most wireless applications, the respective radios are found approximately at ground level, carried by people or cars, or set in homes and offices. For many of these applications, radio signals will be attenuated by buildings, terrain, and the earth’s curvature, which features will mostly control the extent of coverage from a given transmitter. The most useful applications will probably be designed to service the ground-level radios, based on the ground-level attenuation of the existing obstacles. Therefore, it is reasonable that many electrospatial applications will be tied to radio coverage near ground level. Non-overlapping electrospatial geographical/spatial boundaries would naturally be drawn based on such applications.

However, at greater heights above the ground, radio waves encounter fewer obstacles than they do at ground level, which means that they will travel much further before they eventually decrease to the E_0 level which signifies the edge of the occupied electrospacetime region. This means that neighboring electrospacetime regions might have non-overlapping footprints at lower heights, but that they would overlap the neighboring electrospacetime regions at greater heights above ground. This would violate the basic requirement to stay out of the neighboring electrospacetime regions.

Figure 7 shows predicted coverage areas for a single transmitter, where the respective coverage areas are defined by a common field strength, but differ by being measured at different heights above the ground. The large difference in these coverage areas shows that the geographical area in which signals remain below E_0 is quite dependent on the height at which the E_0 limit is enforced. Therefore, it may be useful to express spectrum rights for terrestrial flexible-use licenses with a standardized “statutory measurement height” (SMH). At heights above the SMH the electrospacetime rights would continue to prevent the installation of transmitters by foreign users, but encroachment by signals transmitted by neighboring licensees at levels above E_0 would be allowed.

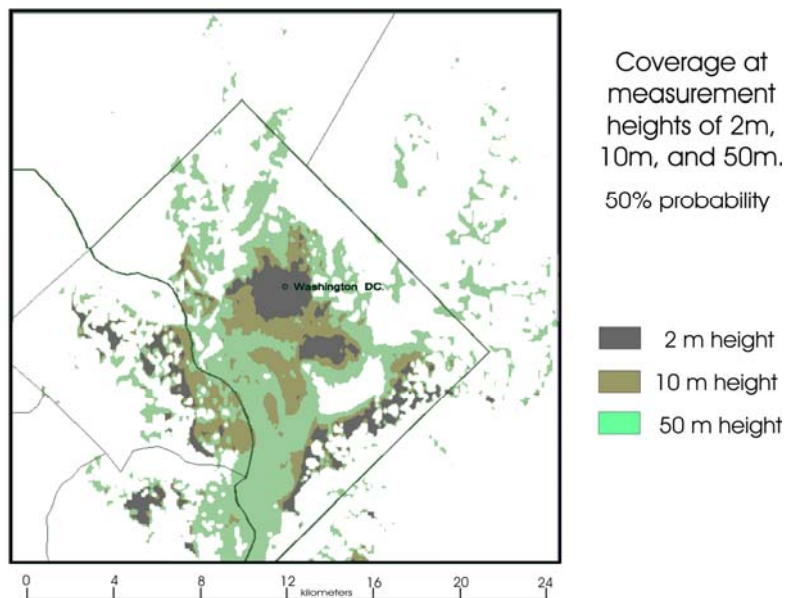


Figure 7 - Coverage as a function of measurement height

If a great majority of radio systems using flexible-use licenses operated with antennas at about the same height above ground, we would presumably choose to match the SMH with that height. However, it is likely that the flexible-use bands will contain a large variety of systems, with antennas deployed over a wide range of heights. Lacking a better criterion, one possible

reasonable value for SMH could be 10 m above ground.²³ This is a compromise between lower heights – where propagation is greatly affected by transient and small structures like houses, cars, vegetation, etc – and higher heights – which are far from the majority of users and where confirming measurements may be more difficult

The use of an SMH could allow the initial distribution of non-overlapping, contiguous electrospacetime regions. However, it would mean that the E_0 limit is not enforceable at heights greater than SMH. Users on the 30th floor of a tall building may receive interference from a neighboring electrospacetime user, but SMH = 10 m means that the E_0 limit is enforceable only at or below a height of 10m. The initial SMH value could be adjusted through side-agreements with neighbors, if necessary.

Of course, one could insist that the electrospacetime regions not intersect at greater heights, but that would greatly restrict the use of the electrospacetime regions on the ground (which is where the terrestrial electrospacetime is especially useful).

Figure 8 shows some of these features. Signals radiated from a transmitter in one licensed electrospacetime region (LER) are shown propagating into an adjacent LER. The LERs have two height features. The top of the LER – possibly set at 30 km – sets a lower limit where foreign transmitters could be located. Above this limit, satellites and other non-terrestrial transmitters

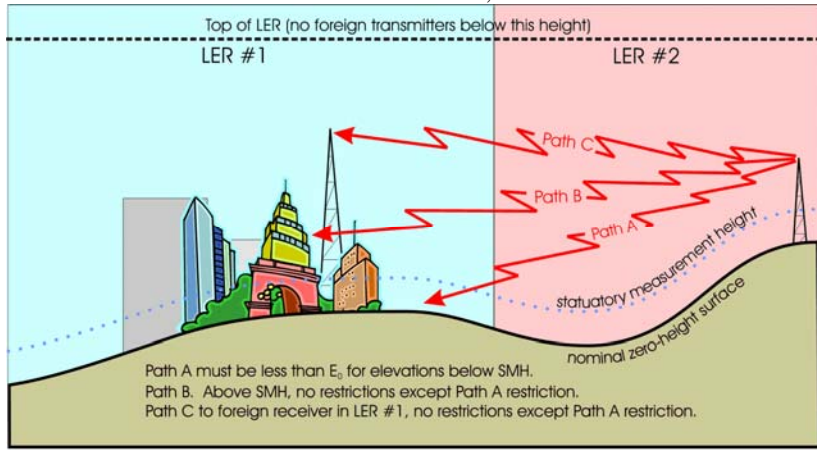


Figure 8 - Flexible-use height considerations

Formatted: Font: Bold

Formatted: Centered

²³ The use of a different height is possible. It may also be important to define exactly what is meant by ground level, especially since modern construction also can greatly change the elevations around a building or overpass, etc.

could be positioned under other licensing arrangements. The statutory measurement height (SMH) is the other height limit, and the E_0 limit must be observed below this height. Therefore, Path A (signals propagating below the SMH limit) must be held to signal strength less than E_0 . Path B shows a signal propagating to locations above the SMH; these signals do not need to be less than E_0 . However, since the Path A signal is constrained in amplitude, the Path B signal will likely be fairly small also. Path C shows a foreign signal transmitted to an operational foreign receiver site, located where the desired foreign signal is strong enough for reliable operation. This is permitted, since there are no limitations on the operation of receivers in foreign LERs.

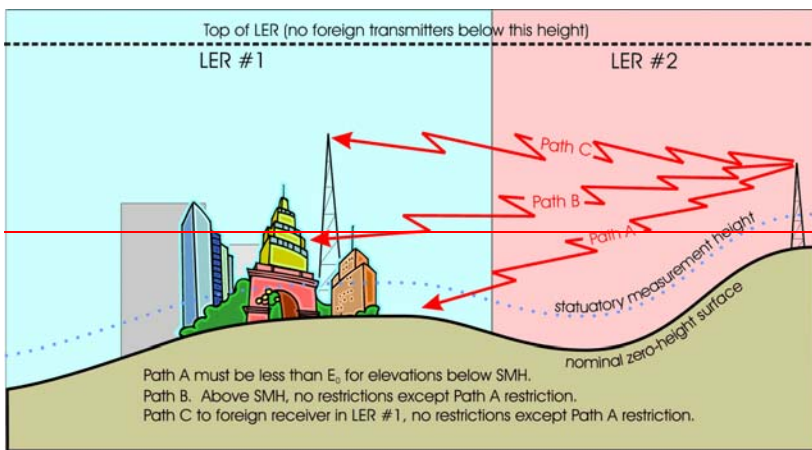


Figure 8—Flexible-use height considerations

The use of a 10 m measurement height specification for an electrospacetime region could lead to problems for wireless users in tall buildings, where many overlapping distant signals might be received at the same frequency. In this case, wireless operators would need to obtain all of the nearby electrospacetime regions at a given frequency in order to control that frequency at greater heights above ground. Similarly, signals transmitted from the top of the building would need to remain within the licensed electrospacetime region, as measured at 10 m above the ground.

If the wireless users were at the top of a tall mountain, instead of tall building, the problems would be somewhat different. Since the mountaintop would be at “ground level” by definition, there would be no exclusion for signals above 10 m height. This means that all of the electrospacetime licensees would need to be careful that their signals did not reach the mountaintop at levels greater than E_0 . Likewise, the mountaintop licensee would need to be careful that signals transmitted from the mountaintop did not reach other electrospacetime regions at signal strengths above E_0 .

5.5 The Probabilistic Nature of Propagation

Figure 3 in Section 3.4 shows one typical way to graph expected signal levels across a

geographical area. The three different colors represent the predicted signal strength (at three different signal levels, separated by 6-dB increments), transmitted by an omnidirectional antenna at a specific height and transmitter power. Other prediction parameters used in generating this map include an assumed height for the receiving antenna and some probability specifications. This section deals with the probability specifications. Figure 9 shows how coverage area changes with different probability specifications.

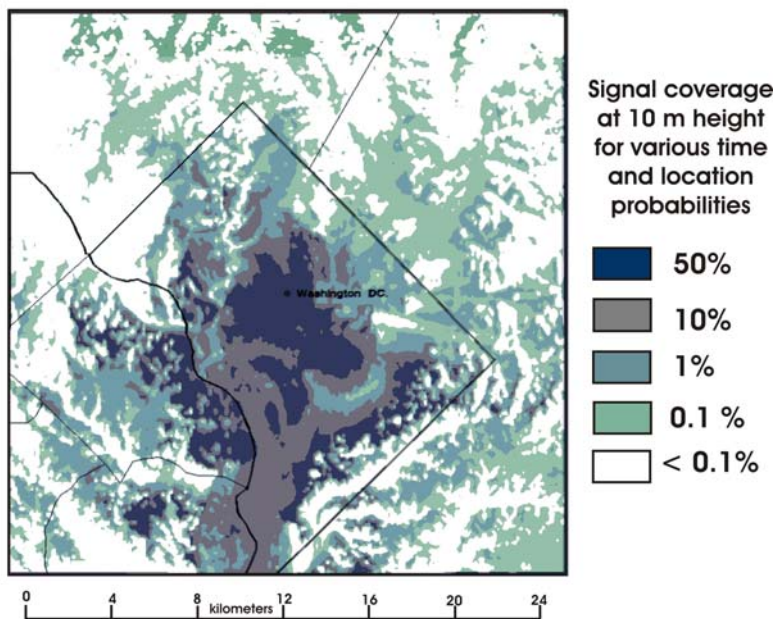


Figure 9 - Coverage as a function of probability

Figure 9 shows propagation as a function of five different probabilities (or reliabilities), respectively $p = 50\%$, 10% , 1% , 0.1% , and $< 0.1\%$. The white areas on the graphs are areas where the signal is received at the selected signal strength for less than 0.1% of the time and locations. Note, that the white areas should not be interpreted as areas where the signal is never received (i.e., probability = 0) at the selected signal strength. The probability specification on most propagation programs includes two separate probabilities: location probability and time probability. The location probability includes the effect of moving around in a general geographical area with certain statistical characteristics. The time probability concerns the changes that occur as a function of time, based on the time-based changes in weather, multipath, nearby moving vehicles, etc. For the purpose of simplicity, we have chosen to graph situations where the time probability and location probability have both been set equal to the indicated probability.

Figure 9 shows that there is an area near the transmitter, where a given signal level is exceeded at least 50% of the time and locations, providing relatively reliable service. However, there are much larger geographical areas where that signal level is occasionally received. These “occasional-reception” areas get much larger as the probability specification gets smaller. Although the ideal flexible-use rule states that the signal should *never* be greater than E_0 outside the LER, propagation modeling shows that occasional higher signals may occur at substantial distances. There is always the remote possibility that an airplane (or a fleet of airplanes, or “atmospheric ducting,” or a meteor shower, or whatever) will fly into exactly the right part of the sky to produce a reflected line-of-sight radio signal that will be greater than E_0 . Of course, we would not expect that this unusual event would happen very often. However, if it ever did happen, should we hold the transmitter owner to be in violation of the flexible-use rules? Should the transmitter owner be subject to criminal penalties or civil damages for a rare or one-time “act-of-god”?

The rule prohibiting signals greater than E_0 outside the LER could be understood to mean more exactly that no areas outside the LER be permitted where the probability of the signal being greater than E_0 is greater than possibly $p = 1\%$. This type of de minimis exemption from violating licensed rights would reasonably constrain the level of signal outside the LER, while allowing productive use inside the LER. The efficient value of p may not be 1%, and to be complete the rules might also have to specify other conditions (including the maximum size of the area over which p is averaged).

In practice, those who wish to argue that their spectrum rights have been violated can present measurement evidence that signal levels higher than E_0 are observed outside the LER at a probability above p . However, individual selected measurements would not necessarily constitute evidence that the statistical bounds were being exceeded, and such measurements would need to be carefully executed and documented to meet a higher burden-of-proof over a statistically-acceptable range of times and locations.

The problem described above can be thought of as an “edge” problem, where the edge of the electrospatial volume of rights is meant to cover propagation that may be ragged geographically and variable temporally. One partial solution to this problem is to consider only larger electrospatial regions, where the ratio of interior space to “edge” is larger. Using geographically larger electrospatial regions, the “edge” problems will affect a relatively smaller percentage of the total space. If the geographical size of the electrospatial region is increased sufficiently, the edge problems will mostly disappear.

6. DOCUMENTING AND ENFORCING SPECTRUM RIGHTS

6.1 Rights Database

Licensed electrospatial regions (LERs) describe the “shape” of the 7-dimension electrospatial volume, within which a user can emit radio signals. One role of government in a flexible rights regime would be to ensure that all LERs and any future modifications are accurately described in an up-to-date database. This database is vital for the government to be able to enforce rights.

Some or all of this data should be public to allow licensees to identify parties who may be violating their rights and to help spectrum market participants identify licensees with which to negotiate.

Assuming parties stay jointly within their electrospacetime rights, there should be no limit to bilateral or multilateral contracts among individual electrospacetime neighbors. Such agreements might allow one party to encroach on the rights of another, for example by emitting signals outside the geographic boundaries at powers over E_0 . An agreement might require that interfering signals be suppressed at heights above the SMH or require suppression at levels different from E_0 . The question arises whether private contracts in spectrum rights should also be included in the public database, especially when the agreement may result in higher levels of interference to a third party.

The question arises as to how the boundaries of the electrospacetime volumes would be expressed in the database. Because these rights may be complex, regulators must give considerable thought to their technical representation. New dynamic spectrum applications may have to query the rights database very quickly, so regulators may face a tradeoff between a very finely scaled map of boundaries and the ease with which one may determine whether a point falls inside or outside the boundaries. A map created by a propagation model (electronically represented) could draw the boundary of geographic rights. Another way might be to approximate (with arbitrary tolerance) the results of a propagation model with a mathematical solid model, which represents 3-D geographic boundaries in a way that allows analysts to ensure easily that adjacent rights are non-overlapping. Sophisticated solid modeling tools exist for manufacturing and other applications, and some allow complex free-form volumes whose surfaces are non-uniform rational B-splines or other piecewise polynomial surfaces. These mathematical approaches would allow clear regulatory delineation of even the most complex propagation predictions. However, one use of these descriptions will be to know whether various boundaries intersect with other boundaries, so the selected models must be relatively easy to search and test for various relationships.

6.2 Rights Protect Against Encroachment, Not Interference

Interference is importantly distinct from encroachment. “Encroachment” means that a signal from a neighboring LER is illegally present at a level higher than E_0 . “Interference” means that an unwanted signal is degrading the performance of an operating radio system. In some cases, encroachment will cause interference (i.e., an unwanted, unauthorized signal causes a receiver to distort or obscure the message in the intended signal). In some cases, encroachment will not cause interference (because the receiver processing is able to reject the encroaching signal, because the desired signal is much stronger than the encroaching signal, or because the desired signal or a victim receiver is not present). In some cases, interference is present without encroachment (because the receiver is not adequate to reject a lawfully present unwanted signal). Under flexible-use rules, the transmitter operator is required to eliminate encroachment, while the receiver owner is responsible for the desired level of interference.

The question arises whether there should be a minimum range in any dimension of a licensed electrospacetime volume. LERs that are geographically too small may not provide an area of usable

signals without encroaching on neighboring LERs, but small LERs may be necessary as part of larger sets of rights in areas where distant “islands” of higher signal strength may propagate.

Recalling the map in Figure 9, the uncertain and variable nature of radio propagation and the strong effects of terrain, structures, and other physical features could make it very difficult for licensees to construct their rights holdings to totally cover all their possible signal propagations. Some researchers have noted that if a licensee does not control all of the small bits and pieces of the rights their system uses, others who own those rights could hold licensees hostage. This potential “encroachment troll” problem would induce licensees to acquire and use their rights holdings very cautiously, including potentially foregoing beneficial uses near the boundaries of their LER. Hatfield and Weiser [2005] discuss the problematic properties of radio propagation signals and rightly note that much of the spectrum policy literature fails to grapple with how such properties can be squared with simple clear rights of access.²⁴ Given that a licensee’s signals might on occasion inadvertently encroach into areas to which they do not hold rights, a challenge arises in crafting a system of rights that distinguishes between trivial encroachment and violation of enforceable rights. Exposing licensees operating in good faith to unexpected and costly litigation is problematic, but so is denying legitimate rights holders a way to redress encroaching signals.

Expressing the E_0 limitation in a statistical sense, as described in 6.5, is one way to help ensure that de minimis emissions above E_0 do not trigger litigation. Licensees could also protect themselves by controlling a sufficiently large geographic area within their electrospace volumes such that they lower the risk of signals spilling into others’ territories. In addition, licensees could enter into private contracts with those upon whom they may potentially encroach so that they can resolve transgressions before the fact and obviate litigation.

In addition to these means by which private actors can minimize the risk of encroachment trolls, it may be useful for regulators to require those who wish to claim their rights are encroached on to show evidence that they are harmed or potentially harmed by the encroachment, similar to showing legal standing for a lawsuit. Precise rules for showing standing must be crafted carefully so as not to blur the system of clear and enforceable rights that is key to a functioning market for spectrum access. Another approach could be to require claimants to go through binding arbitration in which expert arbiters can expeditiously adjudicate rights and mitigate unreasonable hold ups.

6.3 Licensed rights and transmitters

The signals from a licensee’s new or modified radio system must comply with the licensee’s particular set of licensed rights. The question arises how the spectrum regulatory system should ensure that compliance. Should government regulators or expert third parties review the radio system or revisions thereto before harm occurs? Should licensees be able to bring new signals

²⁴ Weiser, Phil and Hatfield, Dale, "Policing the Spectrum Commons" . Fordham Law Review, Vol. 74, No. 2, pp. 663-694, 2005. Available at SSRN: <http://ssrn.com/abstract=704741>

on line without prior approval, to be held accountable after the fact if they trespass on others' rights? In general, new or modified systems should be run through some kind of propagation model to see whether predicted signals would lie within the LERs of the system owner. Key questions for the regulatory system are who should undertake that analysis and what protections the validated results should afford the licensee. In addition, the question arises whether the process should allow an opportunity for public comment and *a priori* objection to the potential new signals. The policy challenge is to establish a system that balances the risks to incumbents of new uses with the benefits of flexible, dynamic spectrum rights. Strong protection from the risk of interference would help prevent costly damage to incumbent systems but could incur delays, add paperwork burdens, and possibly invite gaming by potential market rivals.

In the discussion below, we consider four approaches to striking this tradeoff: a federal operating license, an expert third party certifier, a band manager, and a self-certifying approach. Other approaches are possible, but we choose these examples to illustrate a range of options. In all of these approaches we envision broadly available high-quality propagation analysis tools and the associated terrain and structure data bases. These tools, no matter who would be responsible for using them, would make the required coverage analysis relatively easy, technically neutral, and consistent across users. For example, electrospatial neighbors could use the same modeling tools to predict their own system coverage and to scrutinize their neighbor's coverage. High quality propagation tools should obviate the role of expensive and time-consuming environmental measurements of field strength in the licensing process.

The first approach, establishing federal operating licenses, may provide the strongest protection from interference, but could incur the highest regulatory burden. It would require each licensee to obtain a separate operating license or other certification for each transmitter from the regulatory authorities before operation.²⁵ To allow owners of neighboring LERs to have a limited (but informed) opportunity to raise objections, the process could require disclosure of technical information about the licensee's intended transmitter and a brief public comment period. The disclosure would include sufficient information to assess potential encroachment (such as propagation model results), but exclude confidential business information (such as data rates and service quality). Public technical information about transmitters would hold system operators accountable as they design their systems and make it easier for authorities to identify culprits in encroachment situations. Lacking credible objections, the regulatory process would automatically grant operating rights.

The second approach substitutes third party expert certification in place of a federal operating license. This makes sense if federal authorities have no particular comparative advantage in undertaking the requisite engineering analysis. Australia requires that all proposed transmitters be "certified" by specially-licensed engineers who check the system designs with suitable propagation models to give assure that the proposed signals would not encroach into non-

²⁵ Licensing an individual transmitter before operation results in a two-stage spectrum licensing process that remains quite analogous to the familiar real estate process. Obtaining an LER is similar to obtaining a parcel of land. However, the right to place improvements on the land by building a house or other structures typically requires obtaining a building permit. The building permit assures neighbors and other interested parties that the specific planned improvements are following the applicable regulations.

licensed areas.²⁶ Certification largely protects licensees from challenge by other incumbents and encroachment trolls.

In the third approach, federal authorities would cede spectrum regulatory authority over relatively large geographic and frequency ranges to band managers. Such entities would devolve legal rights to access spectrum to individual users as the band manager sees fit. Band managers could tailor rules to prevent risks of interference and encroachment trolling to the environment in their bands. The relatively large LERs would ensure that most interference complaints will be among users who are mutually associated with the same band manager (instead of between different band managers, where interference control may be legally more complex). However, the question arises exactly what advantage band managers confer relative to federal regulators, who could presumably impose band-specific rules themselves. Band managers may provide more efficient rules if they have better information or are more nimble to respond to changing market or technical conditions.

A fourth approach would require LER owners to test their pending systems with approved propagation models and retain records that show that the predicted signals lie squarely within their LERs. They could do this test themselves or hire a third party. In the event of encroachment disputes, the licensee would have to show the records indicating their compliance.

None of these approaches fully eliminates the risk of encroachment trolls or actual harmful interference. In addition, all of these approaches require possibly costly and expensive propagation analysis, although good analytical tools could reduce these transactions costs. Finally, all of these options raise the question of how to allow for dynamic secondary markets in flexible spectrum rights when different users operate systems with different transmitters and other technical characteristics. Indeed, such dynamic spectrum markets raise a lot of questions that we have not fully addressed here.

The optimal approach might be to adopt different rules ~~depending~~ in different bands or other subdivisions of the electrospace. This would allow users to choose the regulatory environment best suited to their applications and tolerance for risk. However, any efforts to customize regulatory environments should take pains to allow spectrum resources to reach their highest and best uses through market forces and avoid inadvertently creating another inefficient, balkanized spectrum management system.

6.4 Advantages and Disadvantages of Flexible-Use Rights

The preceding sections described a way to articulate clear rights of access to spectrum while reasonably accounting for challenging physical properties of radiowaves. We argue that it would provide for the economically efficient use of spectrum through market forces and would induce an efficient level of interference.

²⁶ “Flexible Radio Spectrum Access: Moving from device-centric to space-centric management.” Michael Whittaker, FuturePace Solutions, 2006.

Although this flexible-use model includes only a very small number of rules and limits, in most cases we have not offered numerical values for these limits. The flexible-use rules do not make radio propagations simpler, and even under flexible-use rules it will still be necessary to make complex and difficult technical trade-offs. However, unlike the current command and control regulatory structure, licensees can immediately make and implement these decisions to optimize their spectrum use, instead of waiting for a problematic distant regulatory process.

In addition to the power rules (E_{\max}) for licensed electrospatial volumes, it may be efficient for regulators to place additional limits in certain frequency bands. For example, most LMR and cellular/PCS services will benefit from duplex band architectures, where base station receiver frequencies are systematically separated from base station transmitter frequencies. In this quite important circumstance, a rigorous separation of transmit and receive frequencies would be very beneficial, but even a single non-conforming frequency could greatly complicate base station implementation. Therefore, although “maximum-flexibility-of-use” remains a key principle, some applications may benefit from some limitations on flexibility (like duplexing rules).

It is likely that a flexible-use environment will have some disadvantages. What appears as “freedom” to one licensee might appear as a “lack of needed guidance and prescribed practices” to another licensee. A higher degree of technical expertise might be required to put a new system in a flexible-use band. A lack of expertise in flexible-use system design might lead to higher levels of interference in a flexible-use band. The lack of narrow standards in the band might mean that a new system would have to be designed to withstand interference from a much wider variety of possible interferers. This might require a more expensive system and more conservative design than would be necessary in a more traditional band (where only one type of interference would usually be present). Licensees must ensure they have the rights to a sufficiently large geographic area. It is likely that flexible-use bands will also change relatively rapidly with time, possibly requiring continuous “patching” to solve new interference problems.

In addition, we have not addressed how flexible rights would evolve from the current system of command and control rules. Would flexible and traditional rights co-exist in a given band for a time? Would a disruptive conversion technique like band-clearing be needed?

Interference resolution is fundamentally different between flexible-use and command and control bands, since the flexible-use model does not protect receivers from interference. Traditional rules give the regulator responsibility to protect performance of the entire radio system (including receivers). When regulators do not prevent interference between compliant licensed users, this greatly complicates the issue of responsibility for resolving interference problems. The flexible-use rights are much clearer. When there is no encroachment, the responsibility to fix interference is placed on the victim receiver. Because of this fundamental difference in determining liability, simply grafting flexible-use freedoms into established command and control bands may not be successful, unless the boundaries of rights are clarified and non-overlapping. Finally, anyone who benefited from the ambiguity of the command and control system of spectrum access could be made worse off when that ambiguity is eliminated.

7. CONCLUSIONS AND SUMMARY

The previous sections have discussed various alternative regulatory models to control the use of the radio spectrum, along with suggesting services and technologies to which different regulatory models would particularly suited (or not suited). The electrospace is one method of describing how radio signals move through the real world. We argue that electrospace volumes, along with power rules, would provide a usable basis for describing signals permitted under a flexible-use spectrum rights system.

With the preceding discussion as background, the next steps will be to develop a more complete proposal for how a flexible-use spectrum rights environment would actually function, specific bands in which it could be applied, and ideas for a process for transforming rights within the current system into a flexible use system. This would involve developing values for E_0 and E_{\max} in specific bands, along with a complete set of rules and procedures to govern licensing, interference/encroachment resolution, secondary markets, and a rights database.