Principles of Flexible-Use Spectrum Rights

Robert J. Matheson

Abstract: **A serious problem with traditional "command & control" spectrum management techniques is that they do not easily accommodate new technologies and new services. This paper describes the necessary principles of flexible-use spectrum rights which may allow a wide variety of spectrum uses in a single general-purpose band. Based on the electrospace description of the radio spectrum, these principles allow general aggregation or division of licensed electrospace regions via secondary markets, providing rules for how regulatory limits change under aggregation or division. These flexible-use principles limit transmitter behaviors that tend to create a more difficult operating environment for receivers, while making receivers responsible for handling any remaining interference. The author shows how flexible-use principles could provide a basis for real-world flexible-use frequency bands.**

Index Terms: **Electrospace, flexible-use, interference, receiver standards, spectrum management, spectrum property rights.**

I. INTRODUCTION

Spectrum management determines what users and services can be provided at specific radio frequencies. In the past, spectrum managers have usually assigned radio licenses that tightly prescribe exactly what frequency, bandwidth, modulation, transmitter power, geographical location, services, and type of user can be associated with a specific radio site. Since a central authority assigns a specific use in each frequency band, radio users can be squeezed together in a band just tightly enough for maximum efficiency, but not too tightly to cause interference. This method of spectrum management—called "command & control"—has been the traditional way to simultaneously ensure high efficiency and freedom from interference.

The command & control technique has a problem in smoothly accommodating new services or new technologies for which no associated frequency bands have been established yet. How can a new service, "A," be offered, when there are no rules and frequencies at which this new service can operate? Equally frustrating, other services, "B" and "C," may have less demand than anticipated, and their corresponding frequency bands will remain mostly unused. Keeping in mind the lack of application flexibility as the major problem of command & control regulations, this paper describes flexible-use rules. These rules allow spectrum to be used for a wide range of user-selected services, and freely traded, aggregated, and divided via a secondary market without regulatory permission.

Manuscript received October 13, 2005.

The author is with the Institute for Telecommunication Sciences of the National Telecommunications and Information Administration of the U. S. Department of Commerce, USA, email: Matheson@its.bldrdoc.gov.

This paper is adapted from "Flexible-use spectrum rights: A tutorial," in *Proc. ISART 2005*, Mar. 1–3, 2005.

The views, opinions, and/or findings contained in this report do not represent an official position of the National Telecommunications and Information Administration or of the U.S. Department of Commerce.

Each of the three major regulatory regimes—command & control, non-licensed, and flexible-use—will have specific sets of applications that fit particularly well within that individual regulatory regime. This paper does not claim that "flexible-use" is broadly superior to other regulatory regimes, but rather that it could accommodate some types of applications better than the other regimes. We envision different sets of regulatory features applied to different frequency bands, each set able to best serve various users and services. The flexible-use principles which we will describe in this paper should be understood as applying only to (currently non-existent) flexible-use frequency bands, without prejudicing in any way the regulatory practices that properly apply to other bands.

Throughout the remainder of this paper, the more general term "electrospace" will be used in place of "spectrum" to eliminate confusion over whether spectrum refers only to the "frequency" dimension or to other dimensions of the radio signal environment as well. The electrospace will be described in more detail in Section II, which will show how the presence of radio signal energy can be described in a specific and unambiguous way using the electrospace model.

Section III will describe a set of rules that allow the establishment of a flexible-use regulatory regime. Frequencies are acquired and used and interference is controlled via the following major features.

- All signals must remain within their respective licensed electrospace region. All signal levels must be less than E_0 outside their electrospace region.
- There are scalable limits on transmitter power or signal strength at ground level.
- Receivers are unregulated, without any guarantee of freedom from interference.
- Unlimited aggregation or division is permitted via secondary markets along all electrospace dimensions.

The crucial factor in establishing flexible-use rules is the degree to which interference can be prevented, while still permitting wide flexibility of use. The proposed rules depend on an understanding of factors that cause interference in receivers, and the rules could change if there were major changes in the performance of receivers.

II. A DESCRIPTION OF THE ELECTROSPACE

The electrospace is a formalized description of the radio signal environment, as it might be seen by a hypothetical ideal receiver at a given electrospace "location" that is defined by the seven electrospace dimensions [1], [2]. It applies to all types of radio systems and all regulatory environments. It is particularly useful in flexible-use applications, where it provides an unambiguous description of licensed electrospace regions.

The electrospace describes radio signals, which means that it describes the domain of transmitters and propagation paths. It

1229-2370/06/$10.00 © 2006 KICS

Table 1. Electrospace dimensions.

Quantity	Units	# of dimensions
Frequency	kHz, MHz, or GHz	1
Time	seconds, hours, or years	1
Spatial location	latitude, longitude, altitude	3
Angle-of-arrival	azimuth, elevation angle	2

omits any description of receivers. Although any real radio system must consider all system components—i.e., the electrospace and receivers—it is appropriate to divide these two components for regulatory purposes. The electrospace describes radio signals that can cause interference to other users—an externalized cost that must be regulated. Receivers, however, cannot cause interference to other users. Therefore, receivers have no associated externalized costs that need to be controlled by regulations. The electrospace model looks only at aspects of spectrum that absolutely require regulation—namely signals—while leaving receivers to the marketplace.

The electrospace describes the radio field strength at a given electrospace "location" that is defined by the 7 electrospace dimensions. These 7 dimensions are all independent of each other, which means that the electrospace can be considered to be a 7-dimensional hyperspace. A "location" in the electrospace can be described by assigning specific values to each variable. It should be noted that different investigators have sometimes included other variables in the electrospace, like polarization and modulation. The set shown in Table 1 is a useful starting point, and little harm is done by omitting some marginal variables such as polarization and modulation.

The physical location of a test point or hypothetical receiver is defined by the three spatial dimensions. The field strength characteristics at that location are described by the remaining variables, including the frequency, time of occurrence, and angle-of-arrival. In a frequency band whose licensing is based on the electrospace, a numerical limit "E_0" will typically be established, such that field levels in excess of E_0 are considered to be signals, which are not permitted outside of the user's licensed electrospace region. An electrospace region consists of all points within a described 7-dimension hyperspace volume. An electrospace region is typically used to denote the hyperspace volume defined by an electrospace license (the licensed electrospace region) or the hyperspace volume occupied by a signal (the region where the field strength is greater than E_0).

An ideal receiver can theoretically separate any radio signals that differ in at least one of their 7 electrospace dimensions. The receiving antenna is considered to be part of the receiver. For example, two co-located radio receivers could function without interference if their desired signals were at different frequencies, or if the signals occurred at different times, or if the signals came from different directions. Radio signals using the same frequency, operating times, and angles-of-arrival could be separated without interference if the receivers were present at different locations.

Frequency: The frequency dimension of the electrospace is a description of the frequency or range of frequencies (bandwidth) of the signal.

Time: The time dimension describes the time of occurrence of the signal. It can be subdivided over a wide range of increments, such as the several-year-duration of a license, regular use of the midnight-to-5 AM time block to update computer data, a one-time use for a 4-hour special events broadcast, or even much shorter times.

Spatial location: The spatial dimensions represent a physical (geographical) location, including altitude, at which the signal is being characterized. A licensed electrospace region includes a volume of physical space, within which licensed signals can be greater than E_0. Unfortunately, there is no practical way to confine radio signals within an arbitrary selected area. Therefore, although one might select a geometrically simple spatial region, the region might be difficult to use efficiently. For a transmitter in hilly terrain, there are many distant locations where a signal is larger than at many closer locations. In addition, a signal that is less than E_0 at ground level will often be larger than E_0 at greater heights above ground. This makes it necessary to describe the height assumptions when describing electrospace regions geographically. Transmitter power, details of the terrain, and directional transmitting antennas are operative in establishing the spatial boundaries of the electrospace region occupied by a given transmitter.

Angle-of-arrival: This factor describes the angle-of-arrival of radio signals at a given location, including the possible effect of multipath components scattered from many objects in many different directions from the receiver location. Note that this factor is usually not created by directional antennas. The pointing direction of transmitting antennas primarily affects the spatial dimensions of the occupied electrospace, i.e., the geographical areas where signals are larger than E_0. Directional receiving antennas exploit existing electrospace direction-of-arrival characteristics, but they do not create them. Recently developed multiple-input multiple-output (MIMO) technology exploits multipath reflections coming from different directions, handled by multiple transmitting and receiving antennas and mathematically processed to generate independent transmission channels. MIMO technology can be considered to use a generalization of the angle-of-arrival dimension of the electrospace.

III. FLEXIBLE-USE RIGHTS IN THE ELECTROSPACE

Section II described how the radio signal environment could be unambiguously described via the electrospace model. The current section shows that it is reasonable to use the electrospace model as a basis for a flexible-use regulatory environment, where electrospace regions can be treated as a commodity and as property. Specifically, this means that

- the use of an electrospace region for a wide range of services is controlled by the license holder, without requiring regulatory approval, and
- electrospace regions are subject to normal property transactions—including the unrestricted ability to buy, sell, or lease, subdivide or aggregate, along any or all electrospace dimensions, via a secondary market, without requiring regulatory approval.

The issues discussed here concern only the rights and obligations of the current user of electrospace regions. We are not concerned with whether the user holds the electrospace region permanently or temporarily, or whether the user gains those rights via an auction, a lottery, a secondary market, or any other mechanism.

A. Ideal Flexible-Use Principles

Flexible-use property rights are based on the single rule that the user cannot radiate any signal outside their licensed electrospace region (including frequency band, geographical area, or authorized time of operation) at a level higher than a small value, E_0. The licensed user can provide any radio services, operate transmitters of any power, at any frequencies, locations, or times within their electrospace region—subject to the rule. Licensed users should assume that within their licensed electrospace regions all signals generated by other users will also be below the same small value, E_0. Presumably, the user will design his own radio system so that it can operate in the presence of small, unwanted foreign signals (smaller than E_0).

Receivers that can adequately reject any signal within another electrospace region are called "ideal receivers" and these proposed rules are called the "ideal flexible-use" rules. As long as all licensed users of non-overlapping electrospace regions follow the above condition, each licensed user will operate without interference—no matter what the other licensed users are doing within their own non-overlapping electrospace regions.

Since there have been no assumptions about the actual sizes of the various users' electrospace regions, there is no reason to suppose that changing the size or number of electrospace regions would cause any additional interference. Therefore, there is no reason to prevent electrospace regions from being divided or aggregated as needed, presumably by whatever arrangements are easiest, including unregulated secondary markets. In summary,

Ideal flexible-use rules (assumes ideal receivers):
1. Transmit without any restrictions inside your licensed electrospace region.
2. Keep your signals below E_0 outside the licensed electrospace region.

Although the electrospace model is critically based on a specified spectral power flux density limit, E_0, which cannot be exceeded outside the licensed region, it is not obvious what numeric value to choose for E_0. E_0 has the units of W/MHz/m^2, and it includes signals 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 usually 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 modulation/signal processing gain. 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.

B. Practical Limitations on the Electrospace

Although the electrospace model is conceptually powerful and potentially very useful, there are a few important problems with its application to the real world. One major problem, nonideal receivers, will be discussed in the next section. Other problems are discussed in this section.

The arbitrary division of the electrospace along any selected dimensions—while theoretically possible—may or may not produce a useful division in the real world. Arbitrary spatial regions, for example, may not match easily achievable propagation/coverage areas. A more marketable spatial division technique may be to use propagation models to determine easily achievable coverage areas and to divide the electrospace regions in a corresponding way. The angle-of-arrival dimensions may be compromised by unintended scattering from the terrain or by lack of sufficiently-narrow-beamwidth receiving antennas (especially at lower frequencies). Division into very narrow time slots may produce systems that are difficult to synchronize properly. Division into very narrow frequency slots may produce unreasonable requirements for frequency stability and Doppler shift, as well as modulation sidebands that spill outside the frequency range. Note that very short time slots are theoretically incompatible with very narrow frequency bands, since the smallest resolvable time, t, requires a bandwidth $\geq 1/t$.

In hilly or mountainous terrain, the areas of spatial coverage may be so discontinuous that it may be difficult to understand how to provide close-in service without causing areas with signal levels greater than E_0 at distant locations. Under these conditions, the model of a coverage area inside some boundary and signals less than E_0 outside that boundary may seem almost meaningless. Raising transmitter power to fill in coverage holes inside the boundary may cause unacceptable signal levels at mountaintops far outside the licensed electrospace region. These circumstances may make the geographical electrospace dimensions difficult to use, especially when smaller geographical regions are considered.

Two points can be raised in defense of the electrospace model. First, the same circumstances also badly affect the traditional command & control models, which typically default to very wasteful worst-case answers. The second point is that only flexible-use rules give enough flexibility and authority to allow detailed local knowledge to make the best use of such difficult situations. Flexible-use rights will allow easier negotiation of local agreements between adjoining electrospace regions, describing disjointed electrospace regions that match the discontinuous coverage areas. Not ideal, perhaps, but better than what could have been salvaged via an adversarial hearing process with a national regulator and clumsy rules.

One obvious application of spatial coordinates is to describe licensed regions using some lines drawn on the ground—e.g., lines described by latitudes and longitudes, a circle centered on a designated location, geographical boundaries, political boundaries, etc. For many applications, radio signals will be attenuated by buildings, terrain, and the earth's curvature, which will control the extent of coverage from a given transmitter. The most useful geographical boundaries might be those drawn by propagation prediction programs, using signal strength at ground level as the electrospace criterion that is legally enforceable. However, almost all signal blockage is greatest at ground level, and raising a receiving antenna farther above ground will usually increase the received signal level. Many radio systems have antennas located on tall buildings, towers, or mountaintops. Therefore, although an electrospace boundary description at ground level may be a useful simplification for many purposes, the success of some antenna tower, airborne, or satellite applications may require complex side-agreements controlling signal strength at various heights above ground.

The frequency dimension can also cause problems. Although a transmitter can radiate any 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 where the signal strength is very high (e.g., close to the transmitting antenna). At such locations, very high signal strength inside the licensed frequency range would need to drop below E_0 immediately outside the licensed frequency range—requiring a very rapid decrease in signal strength over a small change in frequency. Therefore, the requirement that emissions outside a licensed electrospace region be less than E_0 may need to be supplemented by an optional "relative-dB" emission mask that provides a "safe harbor" in locations where the signal strength is very high. This emission mask could be similar to existing transmitter emission masks.

C. Practical Flexible-Use Principles

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

The above scenario is equivalent to saying that ideal receivers can separate any two signals that differ in at least one of their electrospace 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.

A major goal of all frequency management (including flexible-use) is to ensure that a "good-enough" receiver can be built relatively easily and inexpensively. For many applications, the direction-of-arrival dimension is not subdivided, allowing the use of simple omnidirectional antennas. In these cases, an ideal receiver would only need to worry about foreign signals that illegally intrude within the licensed electrospace region and appear at the frequency of the desired signal—so called "in-band" interfering signals. Observing the electrospace rules that control E_0 would be all that is needed to control interference. 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 cause receiver distortions that are seen as interference; this is known as "out-of-band" interference.

Fortunately, radio systems do not usually require ideal receivers for satisfactory operation. Instead, they merely require "good-enough" receivers. A "good-enough" receiver is a receiver whose performance is at least good enough to achieve the desired system performance in the actual radio signal environment. The required performance for a good-enough receiver will vary greatly, depending especially on the presence of very strong signals in the electrospace environment. In general, fewer and weaker strong signals, with greater frequency separation from the receiver tuned frequency, will make it easier to build a "good-enough" receiver.

Practical flexible-use rules supplement the ideal flexible-use rules with a limit on transmitter power, producing a more benign signal environment that allows the use of less-expensive good-enough receivers. The transmitter power limit should be chosen on a principle of maximizing overall benefits—balancing the benefits from less-expensive receivers with the disadvantages of more-restrictive limits on transmitter power. Presumably, different limits could be selected for different bands, maximizing the benefits for various types of systems that could be built in each band. Under these supplemented rules, the important principles that regulate flexible use are now.

Practical flexible-use rules (assumes non-ideal receivers):
1. Transmit within power restrictions inside your licensed electrospace region.
2. Keep your signals below E_0 outside the licensed electrospace region.

Transmitters must follow all of the ideal flexible-use rules, adding a supplemental rule on maximum transmitter power. The limitations on transmitter power will be discussed in more detail in Section III-D, and the way that these limitations scale when electrospace regions are aggregated or divided will be discussed in Section III-F.

A major advantage of these principles (compared to spectrum regulation models that include the receiver) is that the rules are much simpler, leading to much less legal ambiguity about who is responsible for fixing interference situations [3]. Assuming that transmitters obey the practical flexible-use rules, the receiver owner is completely responsible for solving his own interference problems. In an interference situation, the "victim" receiver owner has 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 require compliance.
b. Improve their own system to eliminate the interference. Such changes might involve improving the victim receiver, increasing desired transmitter power (within legal limits), adding better error correction, etc.
c. Tolerate the interference. This includes changing operating procedures, restricting the operation to areas where interference is not a problem, ignoring the problems, issuing the customer a partial refund, etc.

d. Negotiate with the interferer. This negotiation is voluntary for both parties. After investigation of the possible alternatives in (a)–(c), it might appear that an adjustment of the interfering transmitter would be the best way to solve the problem. Negotiations might result in an appropriate business agreement that could become a legal attachment to the respective electrospace licenses.

The opportunity to select an appropriate receiver remains completely with the receiver owner. The practical flexible-use rules help to produce a more-benign signal environment (fewer strong signals) where a good-enough receiver is easier to build. The radio system owner has much better knowledge of his specific requirements than any federal regulator, and the owner is highly motivated to make a correct decision about how to get that required performance. The owner's receiver decisions cause no interference to any other radio system. Therefore, decisions about a good-enough receiver can be left completely in the hands of the radio system owner.

D. A Limit on Power or Maximum Field Strength

The ideal flexible-use rules control external signals at the receiver operating frequency (in-band interference) by requiring that they are always less than E_0 outside of their own electrospace region. The practical flexible-use rules add controls on the presence of strong signals that are radiated legally in other electrospace regions at frequencies different from the receiver tuned frequency. 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 it easier to build good-enough receivers. 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 a much smaller set of circumstances where 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.

Transmitter power alone is not responsible for causing out-of-band interference in receivers; additional factors must also be present. Specifically, out-of-band interference to receivers in public areas results from a combination of three factors.

- High transmitter power,
- transmitter antenna patterns that produce high field strength signals on the ground, and
- the presence of a susceptible receiver in the high field strength areas.

Out-of-band interference could be reduced by controlling the transmitter antenna patterns underneath/nearby the transmitting antenna or by placing transmitting antennas in locations where susceptible receivers will only rarely be found. Instead of controlling the interference solely by limiting transmitter power, it would provide user flexibility to control interference by adjusting transmitter antenna patterns and/or by separating transmitter sites from likely concentrations of susceptible receivers.

Thus, a more effective supplementary rule to protect receivers might include a limit on maximum signal field strength at ground level in public areas. This alternative rule would state that field strength at ground level must be less than $E_{\max}$, where $E_{\max}$ corresponds to a total V/m measured over a large frequency range. 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 have a much wider bandwidth than most transmitters. This limitation must be met only in areas where it is probable 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.[1] 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 unrelated receiver owners.

The value of the parameter $E_{\max}$ is completely determined by practical receiver technology; there is nothing theoretically absolute about this value. If changes in receiver technology cause 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 has made receivers smaller and cheaper, but not necessarily better. Future changes in receiver performance may result from 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 radio (SDR) and cognitive radio (CR) systems 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.

E. Rules for Scaling $\boldsymbol{E_0}$

An important feature of the flexible-use regulatory environment is freedom to aggregate or divide an electrospace region along any or all of the seven electrospace dimensions. This section considers how the limit E_0 for the amount of signal allowed to leak outside of a licensed electrospace region should change when the size of the region is changed. 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

[1]Note 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.

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 electrospace regions comes from an understanding of what constitutes a "signal" that is entitled to leak E_0 into adjoining electrospace regions. The owner of a single region should not be entitled to leak more signal outside his electrospace region by simply claiming that a single electrospace region (and signal) had been divided into multiple electrospace regions (and signals)—each "signal" with a separate allowance for E_0. The exact rules that define a "signal" may be hard to state, since signals are handled in many interesting ways. If ten independent radio signals are combined, amplified by a 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 electrospace 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.

F. Rules for Scaling EIRP and $\boldsymbol{E_{\max}}$

This section describes two possible rules for scaling the power of transmitters under aggregation or division—one for the transmitter power (or preferably, EIRP = transmitter power times antenna gain) case and one for the $E_{\max}$ case. These two cases develop somewhat different sets of rules.

EIRP/transmitter power model: In the case of an electrospace model that includes a maximum transmitter power or EIRP limit, the actual definition is in terms of $Y =$ W/MHz. If a wider bandwidth is divided to give two smaller bandwidths, each of the smaller bandwidths will have a maximum transmitter power proportional to the relative bandwidths of the new regions. Moreover, the maximum transmitter power of the original bandwidth will be equal to the sum of the maximum power of the two smaller bandwidths. This scaling rule for EIRP is a very natural rule for scaling spectrum use rights, since the total allowable transmitter power does not change when a given transmitter is divided into smaller bandwidths or when a transmitter aggregates additional bandwidths.

In terms of scaling power along other electrospace 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 a more powerful transmitter to be employed, simply because the new regional 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.

$\boldsymbol{E_{\max}}$ **model:** An alternative model for scaling transmitter power under aggregation and division is used to control the maximum field strength, $E_{\max}$, at ground level where receivers will be present. 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 specific maximum field strength. Presumably, the cumulative power from multiple transmitters 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. No other constraints are imposed to prevent out-of-band interference from other transmitters.

IV. SUMMARY AND COMMENTS

The preceding sections have described a possible approach to flexible-use spectrum management. This set of concepts could provide a market-based method of acquiring and using frequencies that would give great flexibility in the use of spectrum, while controlling most interference. In the few circumstances where interference might result, the model includes clear rules for assigning responsibility to mitigate the interference. Many spectrum managers believe that access to unused or underused spectrum is the major problem in spectrum management today. Therefore, a spectrum management system that can support almost any type of service and allow users to immediately obtain spectrum via secondary markets could be very useful.

Although this flexible-use model includes only a very small number of rules and limits, numerical values for these limits have not been chosen yet. There are still areas of ambiguity, but some of these can be resolved with "safe harbor" practices or more detailed rules. Many of these ambiguities refer to situations that are also substantial ambiguities under current command & control spectrum management practices. The flexible-use rules do not make the radio world simpler than it now is, and even under flexible-use rules it will still be necessary to make complex and difficult technical trade-offs. However, unlike the current command & control management, the spectrum user is authorized to immediately make and implement these decisions, instead of waiting for a problematic distant regulatory process.

Some limitations on flexible use might be beneficial in certain frequency bands. For example, most LMR and cellular/PCS services benefit from duplex band architectures, where base station receiver frequencies are systematically separated from base station transmitter frequencies. Therefore, although "maximum-flexibility-of-use" remains a key principle, some applications may benefit from some limitations on flexibility.

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. The possible 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 a 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 interferer would usually be present). Problematic propagation in urban or mountainous areas might make it difficult to use small- or medium-size geographical licensed areas. It is likely that flexible-use bands will also change relatively rapidly with time, possibly requiring continuous "patching" to solve new interference problems.

All of this suggests that a traditional single-service frequency band might remain the best place to find a home for many new or old radio systems. However, it is also expected that technological obsolescence may create situations where many new applications no longer fit any of the existing, gradually emptying, single-service bands. A major question is whether additional uses can be painlessly grafted into existing band allocations—possibly using some of the principles of flexible-use described here—or whether some other more disruptive conversion technique will be needed.

Interference resolution is fundamentally different between flexible-use and command & control bands, since the flexible-use model does not protect receivers from interference. Traditional command & control models often take responsibility for performance of the entire radio system (including receivers). The added responsibility for receiver performance makes these command & control models more complex, and it greatly complicates the issue of responsibility for resolving interference problems. The flexible-use model reduces the determination of responsibility for solving interference problems to a simple measurement of field strength. In contrast, command & control regulations often presume that interference from the operation of a new transmitter is mainly the responsibility of the new transmitter to fix, even though the new transmitter is operating within all applicable rules. Flexible-use principles rigorously state that the responsibility to fix this interference belongs to the victim receiver. Therefore, simply grafting flexible-use principles into established command & control bands may not be successful, unless changes are also made to the rules for resolving interference.

Although this paper has described some of the principles of flexible-use bands, it should be noted that many details have not been optimized. These details include especially the specific values for E_0, Y, and $E_{\max}$ that might be chosen for a specific flexible-use band to best match specific classes of technologies or services.

REFERENCES

[1] W. R. Hinchman, "Use and management of electrospace: A new concept of the radio resource," in *Proc. IEEE ICC'69*, 1969.

[2] R. J. Matheson, "The electrospace model as a frequency management tool," in *Proc. ISART 2003*, Mar. 4–7, 2003.

[3] D. Ewing and L. Berry, "Metrics for spectrum-space usage," *Office of Telecommun. Report OT 73-24*, Nov. 1973.

Robert J. Matheson was born in the USA in 1939 and received a B.A. in physics in 1961 and an M.S. in electrical engineering in 1968, both from the University of Colorado. He began work in 1957 with the Institute for Telecommunication Sciences (ITS), the research arm of the National Telecommunications and Information Administration (NTIA). Early work at ITS included measurement of radio noise and development of automated spectrum monitoring systems. From 1988 to 1991, Mr. Matheson directed the work of the Spectrum Division of ITS. Since 1991, Mr. Matheson has been working on technology forecasting, projections of spectrum requirements, characterization of ultra-wideband (UWB) signals, market-based spectrum management theory, and improving the spectrum efficiency of federal systems.