



2005 ISART Tutorial

Flexible-Use Spectrum Rights

Robert J. Matheson
NTIA/ITS.M

matheson@its.bldrdoc.gov

Does not necessarily represent official policies of NTIA



Spectrum Management

- Spectrum management is the process of arranging the use of the spectrum for communications and sensing.

Who gets to use spectrum?

For what functions?

Under what restrictions?

What are the processes to decide?



Outline

1. The Physical World /Engineering

- Electospace
- Receivers
- Interference

- - - 10-minute break - - -

2. Spectrum Management

- General comments
- Command-and-control

3. Flexible Use Rights

- Pure electospace rules (ideal receivers)
- Modified electospace rules (real receivers)
- Extended rules for aggregation and division
- Miscellaneous loose ends



Radio System Basics

Electrospac

Receivers



Radio System Basics - 1

Radio systems have:

- Transmitter (including transmitting ant)
- Propagation path
- Receiver (including receiving antenna)

Communications –
move data from transmitter to receiver.

Sensing –
Compare received signal with transmitted signal to study path (radars, sensors, etc.)



Radio System Basics - 2

- The Electrospace describes the appearance of radio signals: transmitters & propagation.
- A receiver processes the electrospace to receive communications.
- One major goal of frequency management is to arrange the signals in the electrospace so that the desired signal can be separated out by a simple and inexpensive receiver.



The Electrospace

Electrospace (described by Hinchman in 1969) is a 7-variable description of EM field strength (hyperspace).

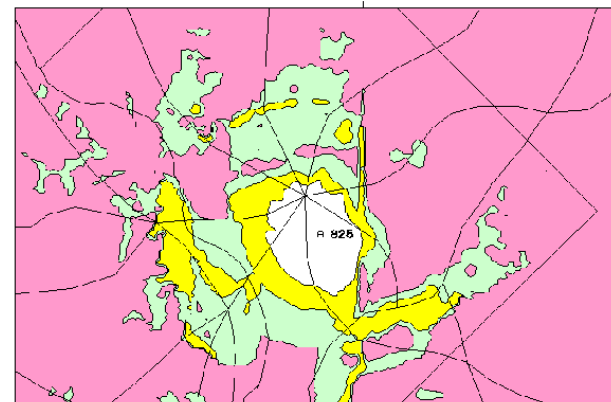
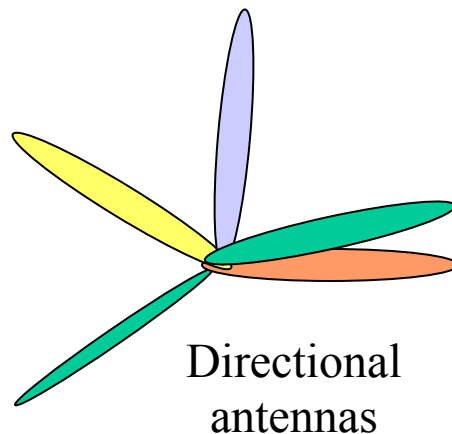
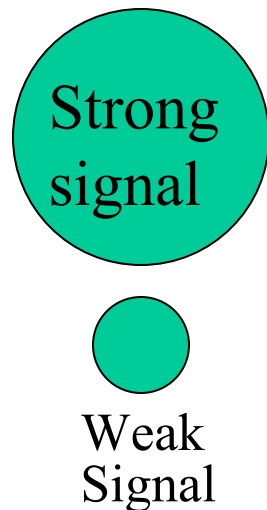
- Physical location – lat., long., altitude 3-dim
- Frequency – MHz 1-dim
- Time - μ S, hours, or years 1-dim
- Direction-of-arrival – azimuth, elevation 2-dim

Electrospace describes radio signals, as they appear from an ideal receiver at the physical location. Note: receiver includes receiving antennas.



Spatial Dimensions

Any spatial region: microcell to BTA, modify with directional antennas. Caution: some areas cannot be used well; coverage affected by terrain and buildings, height above ground. Match spatial to coverage.



825 MHz coverage

Best Transmission (dBm)

Less than 110,000	Area: 10 sq km
	Population: 14000
	Households: 2000
110,000 to 111,000	Area: 20 sq km
	Population: 50000
	Households: 21000
112,000 to 113,000	Area: 40 sq km
	Population: 160000
	Households: 71000
Greater than 113,000	Area: 14 sq km
	Population: 57000
	Households: 23700



Frequency Dimension

The frequency dimension is well behaved and intuitive, with a frequency band often divided up into many identical non-overlapping channels.



Time Dimension

FCC recently mentioned “time” as one additional dimension that they would allow users to divide.

Months – seasonal uses

Hours – to broadcast special football game

Hours – midnight-to-5am daily to send
low cost computer updates.

10 ms TDMA timeslots every 50 ms.

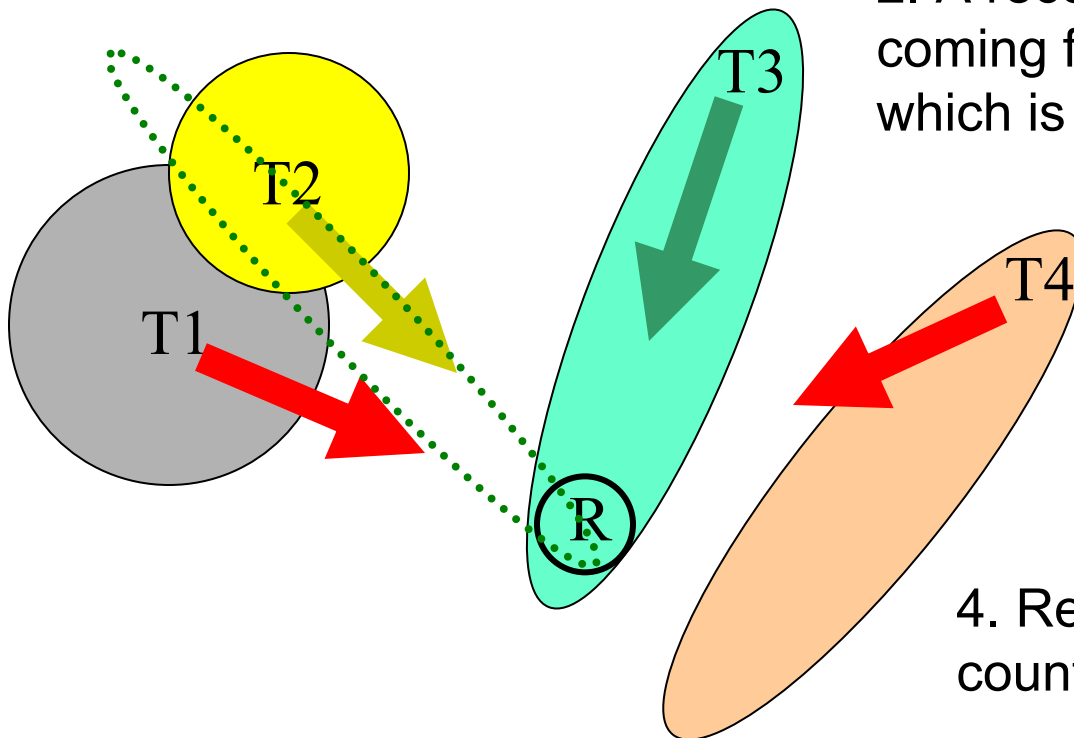
Useful for serving many intermittent activities.



Angle-of-Arrival Dimension

1. Angle-of-arrival is different from coverage area. Consider coverage areas of 4 xmtrs.

2. A receiver will see these 4 signals coming from directions of 4 arrows, which is angle-of-arrival for 4 xmtrs.



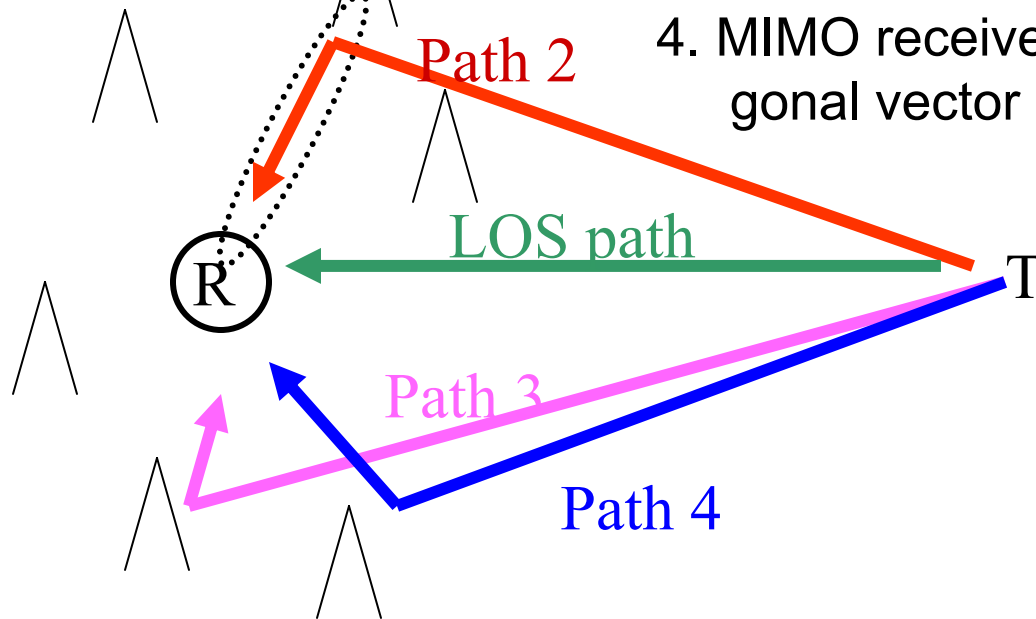
3. A narrow beamwidth receiving antenna can select desired xmtr.

4. Receiver antenna beamwidth counts, not transmitter beamwidth.



Angle of arrival - 2

1. Narrowbeam line-of-sight path (pt-pt microwave).
2. Multiple narrowbeam (NB) paths scattered from landscape.
3. NB receiver antenna isolates each path (difficult).



4. MIMO receiver forms multiple orthogonal vector sums to mathematically generate paths (easier).

5. Future application of generalized angle-of-arrival, using less-expensive techniques



The Electrospace

Electrospace is a 7-variable description of EM field strength (hyperspace).

- Physical location – lat., long., altitude 3-dim
- Frequency – MHz 1-dim
- Time - μ S, hours, or years 1-dim
- Direction-of-arrival – azimuth, elevation 2-dim

▪ A given “region” of electrospace is considered “occupied” if the field strength exceeds “X” W/MHz/m².



Same spectrum, more users

- The electrospace formalizes additional ways to divide up spectrum among more users.
- Point-to-point microwave uses shaped pencil beam coverage areas and divides up angle-of-arrival.
- Trunked radio systems adaptively divide time.
- MIMO mathematically finds multiple paths by processing angle-of-arrival part of electrospace.
- Advanced cognitive radios will be even better at dividing up the electrospace.



Radio Systems

- Receivers and the electrospace create a radio system.
- **Receivers** process the electrospace to give service (if successful) or interference (if not successful).
- A **sufficiently good** receiver can separate any signals having different electrospace descriptions. Interference is caused only when a receiver is not “good-enough”. Might require adaptive antennas to null interference.
- “Interference protection” really means “able to use a cheaper receiver.”



Interference - 1

Interference – any distortion of the processed desired signal caused by unwanted extraneous radio signals (excludes multipath, internal noise).

Interference can be caused by co-channel operation, excessive sidebands, intermodulation in receiver, receiver overload, etc. Note: not by lack of signal.

No sharp line between acceptable and unacceptable interference. All interference is unwanted. Even the possibility of interference is unwanted, since it requires more robust system design.



Interference - 2

- A faulty transmitter causes interference to other users – an “externalized cost” that the transmitter owner has no motivation to control. Therefore, regulations may be needed to control transmitters.
- Externalized cost is either cost of interference, or cost of better receivers to prevent interference. Interference rules establish expected receiver capabilities.
- A faulty receiver causes interference only to receiver user, who is well-motivated to fix receiver. No external controls are needed.



Interference – 3

Interference is **always** caused by an inadequate receiver and could be fixed by a “good-enough” receiver (though “good-enough” for some situations might require adaptive antennas to null out interference, or other very complex and expensive components).

Using better receivers would decrease interference, and/or allow more signals to be transmitted before interference occurred.

Therefore, using better receivers is always expected to improve spectrum efficiency.



Better Receivers

“Better” receiver means anything that helps a receiver more successfully receive wanted signals.

- Better IF filters to remove signals on adjacent channels
- Better dynamic range to reduce IM and overload from strong signals
- Better directional antennas to isolate desired signal from others
- Better intelligence to adaptively re-tune to less crowded frequency
- Better intelligence to change modulation and power
- Etc. e.g., cognitive radios



Receiver standards?

Using better receivers could improve spectrum efficiency.

Therefore: **Require minimum receiver performance standards?**

Yes, but ...

A major objective of good spectrum management is to develop rules that allow the use of cheap equipment (receivers), e.g. :

Keep low power and high power bands separated,

Use large duplex band structures,

Limit maximum transmitter power, etc.

Such features are intended to allow the use of cheaper receivers.

“A major goal of spectrum management is to make the world safe for cheaper receivers.”



Receiver standards

User groups who mutually depend on the performance of other members' radios have good reasons to regulate minimum receiver performance of other members' radios.

Procurement of equipment often simplified by referring to receiver procurement standards.

Receiver standards valuable to **educate user** about typical receiver performance requirements.

None of above primarily concern spectrum management.



Receiver Summary

- Receivers are a vital component of any radio system, greatly affecting performance of system. Receiver performance is an important part of system design.
- All interference is caused when receiver performance is not sufficient for the given electrospace environment. Better receivers can improve spectrum efficiency, at a cost. Poor spectrum management rules may make it necessary to use better receivers.
- The receiver user is usually well-motivated to get a “good-enough” receiver.



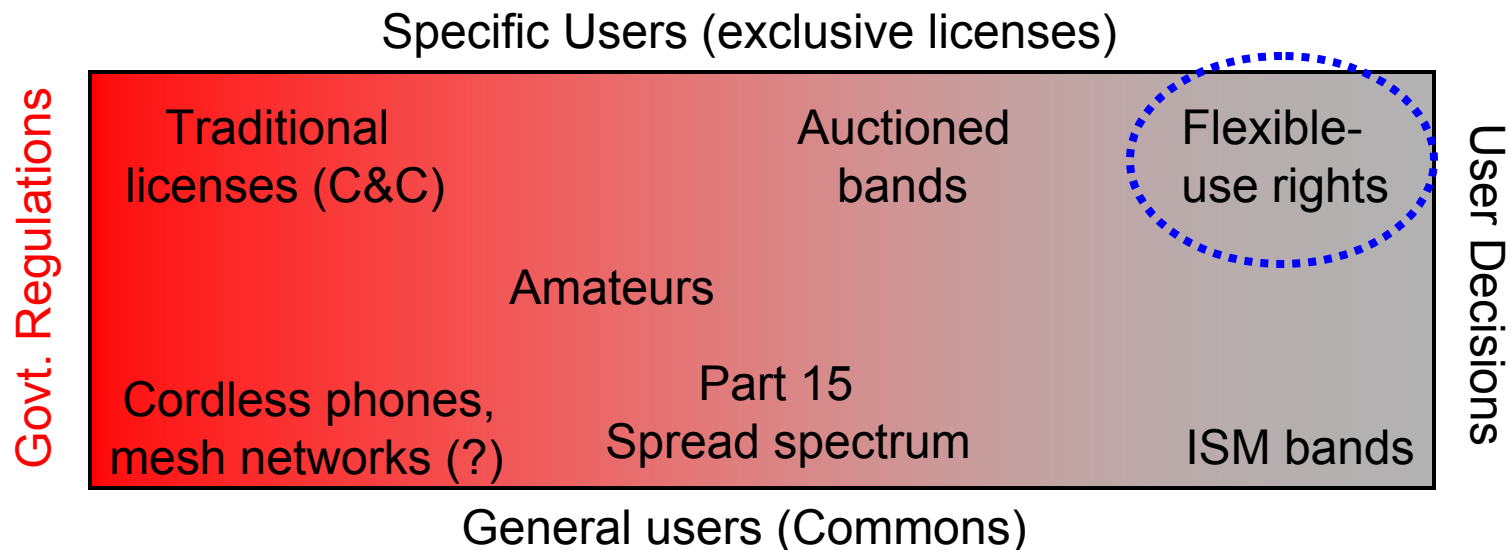
- - - 5-minute break - - -

5-minute break

Next: Regulatory Concepts and
flexible-use rules



Map of Spectrum Rights

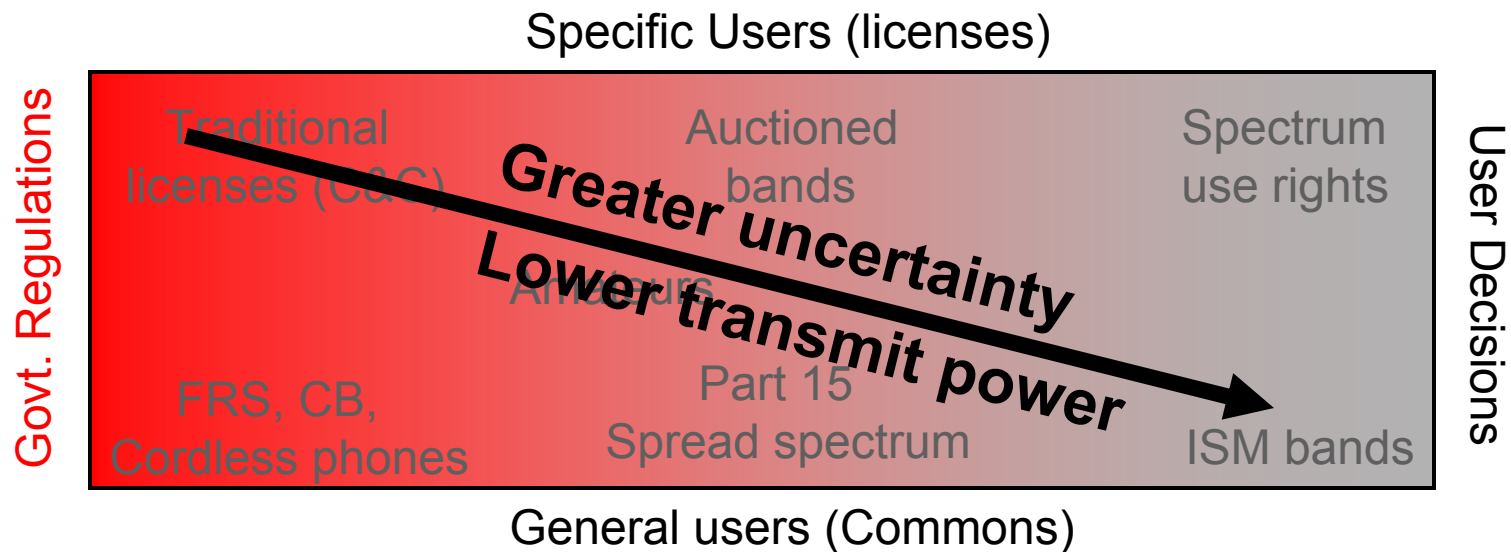


Spectrum management represented on a 2-dimensional continuum.

- Horizontal axis shows decision-making rights, from all-Govt to all-user.
- Vertical axis shows how much a frequency is preferentially given to specific users versus general users.
- Many areas of model are currently in widespread productive use.



Map of Spectrum Rights



People have often thought about radio licenses on an axis between upper-left and lower-right corners.

- Licensed systems with high-power transmitters (traditional C&C)
- Non-licensed systems with low-power transmitters (Part 15)



Complementary Concepts

- Multiple frequency management concepts are complementary, not antagonistic.
- Will always be a requirement for multiple spectrum regulatory environments, because of different system technical and business requirements.
- Real estate analogy - different kinds of property with different rules.
 - Publicly-owned public spaces (parks, highways)
 - Publicly-owned private spaces (classrooms, hospital rooms)
 - Privately-owned public spaces (stores, amusement parks)
 - Privately-owned private spaces (houses, hotel rooms).



Choose the best fit

- Specific services, technologies, and frequencies will often fit much better under one concept than another.
- Actual band rules often mix rules from multiple concepts. Few pure concept bands. Figure out which mix of rules is best for each intended service and technology.
- Easy to apply different rules to different bands. No special advantages to having same rules for all bands, but many disadvantages.



Regulatory Concepts

Command and Control



Command and Control

Regulator makes all of the decisions

- Which bands are used for which services, who is eligible to use band.
- Completely blended engineering and regulation. Definitions for transmitter sidebands, receiver off-frequency rejection, etc.
- Complete band recipe: Service provided, technical parameters, base station sites, service area, frequency re-use distance, def. of harmful interference, etc.
- Guaranteed service, if you follow the recipe.



Band Allocations

- C&C allocates specific bands for specific services. Each band is designed for services, channelization, service area, frequency re-use distance, modulation, receiver specs, allowable users, transmitter power, etc.
- Band allocations include: Mobile, broadcasting (AM, FM, TV), point-to-point microwave, ISM, satellite, PCS, Radar, paging, MMDS, LMDS, etc.
- Each band allocation contains detailed, specialized rules for successful operations in that band.



C&C Receivers

- Each band is engineered for all aspects of operation, including the desired receiver specifications.
- If you operate a receiver meeting the allocated band receiver specifications, you should not get interference.
- If you do get interference, there must be something wrong with receiver or transmitters. Possible need for a (new) transmitter to change operations, even though transmitter meets all nominal specs.



In favor of C&C

- Well-engineered and optimized services, at a given point in time, technology, and social needs. Standardized, stable, efficient, assuming that regulations can keep up with change.
- Highly-differentiated services – mobile, microwave, satellite, radar, broadcasting, Part 15, PCS/cellular, etc.



Problems with C&C

- Slow to provide band allocations for new technology and changing social needs. No bands for new services. Lots of bands for aging, obsolete services.
- Consensus mode of public review slows down any detailed planning. Tend to design conservative, worst-case, less-productive systems.
- Almost impossible to keep up with technological changes in many rapidly-changing systems.
- How to overcome disadvantages of C & C regulations?



Regulatory Concepts

Flexible Spectrum Use Rights



Flexible Spectrum Use Rights

Flexible-use spectrum rights describe permissible ways to transmit radio signals, such that:

1. The rights to use spectrum can be bought and sold on a secondary market, including the ability to divide and aggregate spectrum rights.
2. The user has great flexibility to provide a wide range of services without asking the permission of regulators.
3. The user has a reasonable expectation of operating without interference from other users.

Note: Spectrum use rights (in this context) do not concern how one initially obtains these rights (license, auction, etc.), or whether possession is temporary or permanent.



Pure Electrospace Use Rights

Is there a way to regulate spectrum, such that simple rules allow a wide range of uses, without causing interference to other users?

- **Electrospace** defines how spectrum could be unambiguously described, licensed, used.
- A **ideal receiver** will separate any signals having different Electrospace descriptors.



The Electrospace

Electrospace is a 7-variable description of EM field strength (hyperspace).

- Physical location – lat., long., altitude 3-dim
- Frequency – MHz 1-dim
- Time - μ S, hours, or years 1-dim
- Direction-of-arrival – azimuth, elevation 2-dim

▪
A ideal receiver will separate any signals having different electrospace descriptions.



Pure Electrospace Rules

Only two rules:

1. Keep your signals within your licensed electrospace region. (no signals permitted outside region)
2. Use ideal receivers to avoid interference from electrospace neighbors.

Any services, technologies, architectures, transmitter powers, modulations, etc. are permitted, as long as these two rules are obeyed. (replaces all of 47CFR and NTIA Manual)



Division and Aggregation

Electrospacetime regions can be freely divided or combined along any combination of electrospacetime dimensions, using unregulated secondary markets.

Approach: Electrospacetime description uniquely describes specific arbitrary regions where signals are allowed, using orthogonal coordinate systems. Therefore, no problems anticipated in dividing or aggregating multiple electrospacetime regions in whatever arbitrary ways seem useful.



Perfection, except ...

The pure electrospace rules seem to be the perfect solution to the C&C problems.

But...

1. I can't reduce signals to "zero" outside licensed electrospace region (frequency or geography).
2. I can't buy an ideal receiver.
3. (Otherwise, everything is fine.)

Are there some reasonable approximations that still leave us with most of the advantages of pure Electrospace rules?



Modified Electrospace Rules

Rule 1: Stay within licensed electrospace region.

Define a very small signal, X , as being close enough to “zero.”
 X = “minimum” signal or “allowable leakage” signal.

X is power spectral density (PSD) = $W/m^2/MHz$. Anything above X is defined as a “signal”. All “signals” must remain within the licensed electrospace region (frequency, location, time, direction-of-arrival).

Choose X to be small enough that it would not ordinarily cause interference to normal system performance. X can have different values in different frequency bands.



X = minimum “signal”

- **X is a parameter with two roles:**

For transmitter: X = maximum power spectral density (PSD) allowed outside the licensed electrospace region. Design your transmitters so that your signal is always below X outside licensed electrospace region.

For receiver: X is guaranteed maximum unwanted signal PSD “leaking” from other users at your desired frequency of operation. (X from each user. Make X small enough, so that leakage from several users will not cumulatively cause interference.)



X - Tradeoffs

Smaller X requires better filtering of transmitters outside licensed electrospace region, lower transmitter power, larger distance from edge of frequency range and geographical borders.

Smaller X means smaller interfering signals from other users, less-expensive receivers.

Where is “sweet spot” in the trade-offs?

Differences between X's in various bands is possible. Might be one of the ways that different flexible-use bands are “optimized” for different classes of applications.



Ideal Receivers?

Rule 2: Use an ideal receiver. Pure electrospace rules are only guaranteed to work with ideal receivers that can separate signals perfectly along all electrospace dimensions.

Note that an “ideal” receiver is not needed for interference-free performance, but only a “good-enough” receiver is needed. In many cases, a “good-enough” receiver will be quite simple.

How can electrospace rules be modified to ensure that a “good-enough” receiver is always relatively simple?

Consider how interference is caused in receivers and put reasonable limits on those cases. Limits could be different in various bands.



Causes of interference -1

Co-channel interference (unwanted signals at same frequency as desired signal).

1. Intentionally-radiated co-channel signal, from transmitters using same frequency at distant locations, or
2. Unintentionally-radiated sidebands/spurious signals from nearby transmitters operating at other frequencies.
3. If permitted, intentionally-radiated low-power underlay signals (Part 15).

Controlling mechanism (pure electrospac rules):

Cannot be more than X dBm/MHz/m² interference from each source outside the licensed electrospac. Falls off rapidly with distance from transmitter.



Causes of Interference - 2

Off-frequency interference - Interference caused by **strong signals** outside of intended receiver bandpass, presumably **outside of user's electrospace region**.

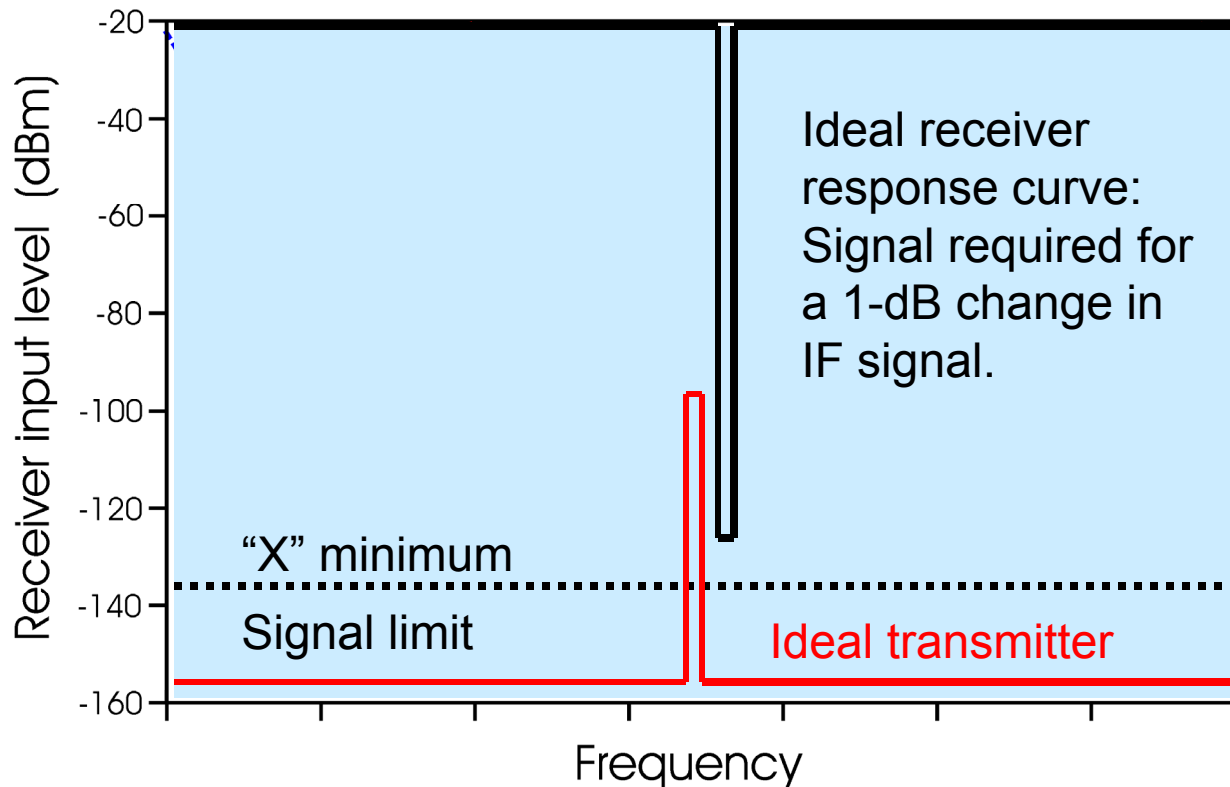
Non-linear: Intermodulation products, desensitization.

Linear: bandpass filter feed-through (worse for close-in strong signals), shielding leakage, image frequencies, LO sidebands and spurs.

A signal environment with strong unwanted signals requires a **better-quality receiver** (higher cost, larger, heavier, greater complexity, more power consumption, etc.) Strong unwanted signals are expensive for receiver owners (just like co-channel interference). Therefore, should try to control strong signals.

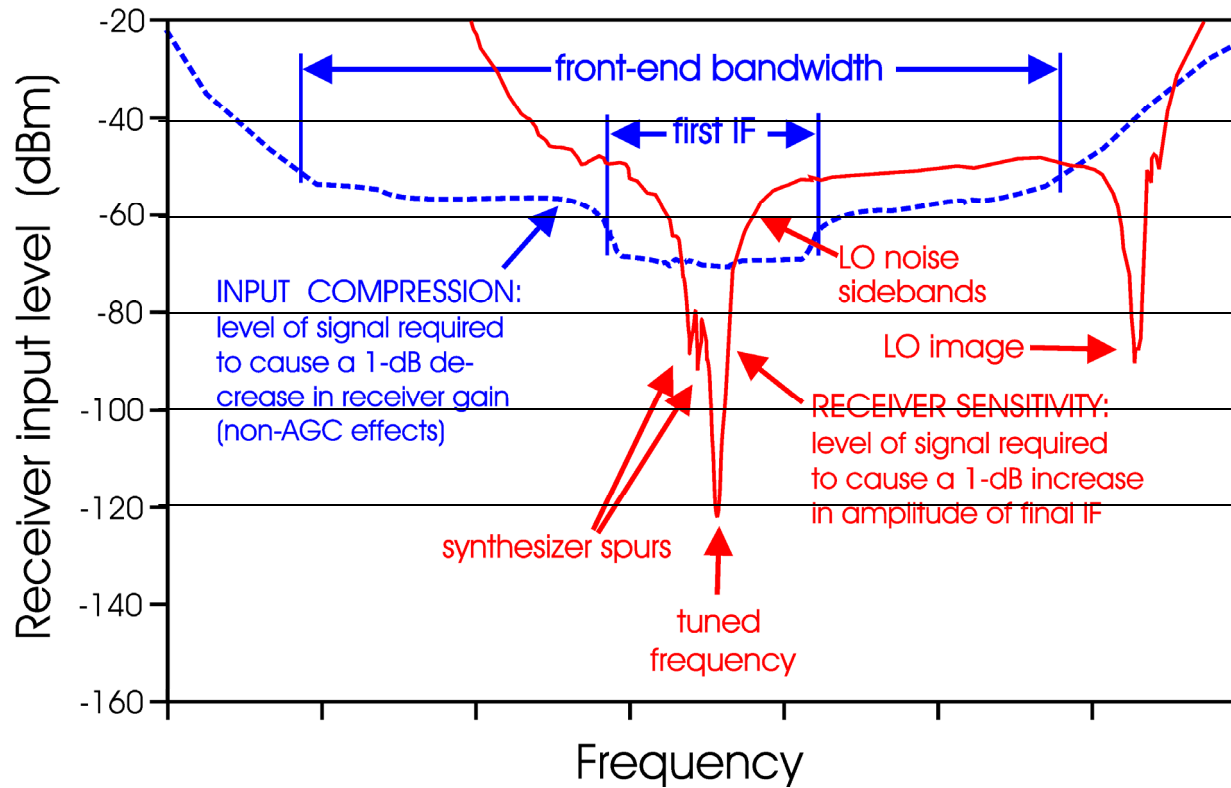


Ideal Xmtrs and Receivers



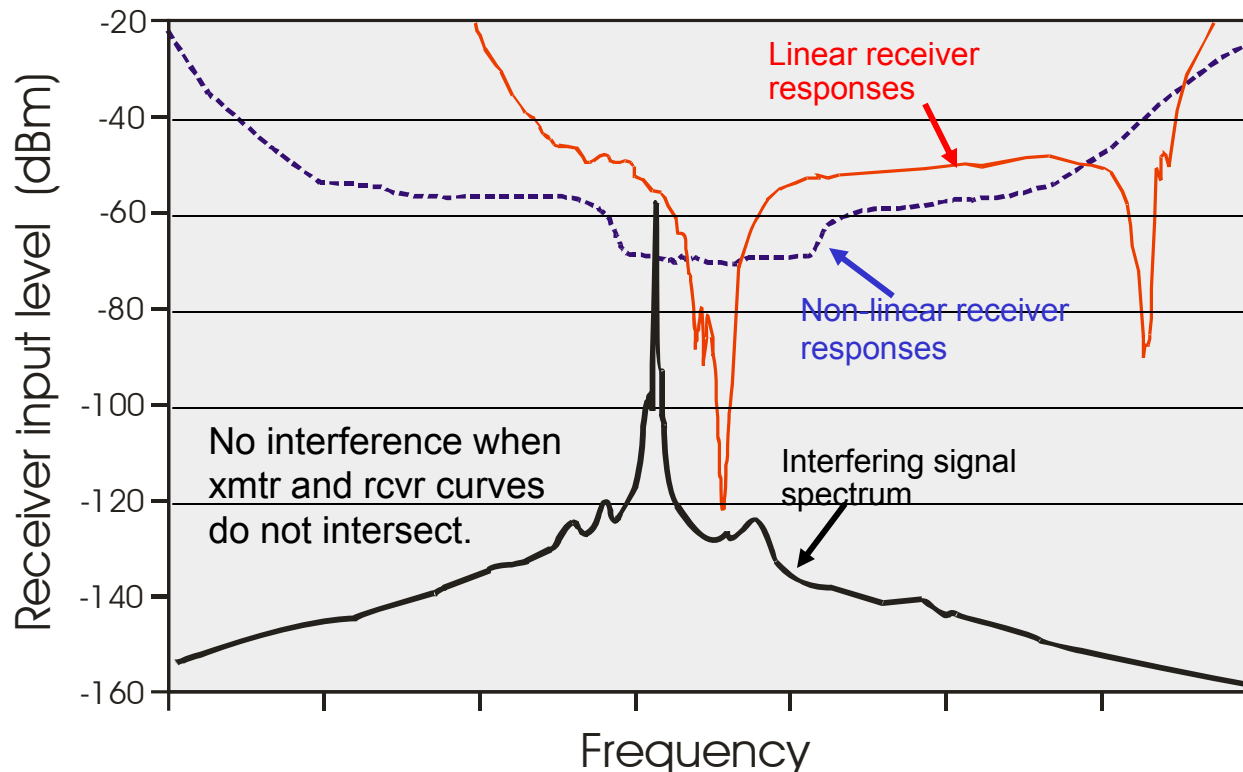


Real Receivers





Real Receivers and Xmtrs



Small signals cause interference only when interfering signal is at receiver tuned frequency. Large signals cause interference at many frequencies. Note: Small signals act like ideal electrospace model.



Control of Strong Signals

Two possible methods to control strong off-frequency signals:

- Control maximum **field-strength** near receivers,
 - - - or - - -
- Control maximum **transmitter power**.

Either – or both – limit(s) could be used.

No limit on number of transmitters, sites, etc. Limit only on maximum power at a location, which aggravates receiver performance requirements.

Completely technology-dependent. High-performance cheap receivers (e.g., receivers with cheap superconducting tracking filters) could greatly change limits.



Field Strength Limit -1

E_{\max} = maximum field strength

- Field strength greater than E_{\max} is not allowed. Rule is the only limitation on the completely flexible use of frequencies; it is needed to control off-frequency interference. (Note: Lack of this rule created the problems in 800-MHz band).
- No limit on EIRP or transmitter power, as long as strong fields don't reach public locations (or cause signals outside licensed electrospace).



Field Strength Limit -2

Field-strength limit only applies in public areas where uncontrolled radio receivers will be commonly found, e.g., ground level underneath urban radio transmitting towers.

Transmitters can broadcast as much power as wanted, but areas of high field strength can cause interference to other radios – an externalized cost to other radio users – which must be controlled.

No limit on transmitter power, but more powerful transmitters will require better control of vertical antenna patterns.

Frequencies near to bands having many battery-powered portable receivers are particularly important to control.



Transmitter Power Limits

Transmitter Power Limits are an indirect alternative method to partially control maximum field strength.

- Licenses specify maximum power = Y Watts / MHz
- Maximum power is proportional to bandwidth. E.g. - adding two adjacent licenses doubles Bandwidth and doubles power.
- Transmitter power limits are not as effective as field strength limits, but they are simpler to manage and they have a nice “feel” to them. (However, this rule would not have prevented 800 MHz interference.)
- A little surprising that this “obvious” rule to prevent off-frequency interference actually does not work very well by itself.



Changes in X , E_{\max} , Y

How should modified electrospacetime limits change under aggregation and division?

- $X = W/\text{MHz}/\text{m}^2$ is already proportional to bandwidth, so allowable leakage power scales linearly with bandwidth of electrospacetime region.
- E_{MAX} is an absolute value. Does not change (receiver overload is not affected by bandwidth of overload signal).
- $Y = w/\text{MHz}$ transmitter power is already proportional to bandwidth, so allowable power scales linearly with bandwidth of electrospacetime region.



Miscellaneous Loose Ends

A few extra rules are needed to prevent “misuse” of the general modified electrospacetime rules by clever engineers.



Division and Aggregation

Primary Rule: Electrospace regions can be freely divided or combined along any dimension. However, interference conditions must not be made worse by any transactions.

Approach: Identify factors that make interference worse (i.e., require better receivers), and adopt rules limiting these factors.

Does not guarantee to prevent interference, but only says that total interference after transaction should not be worse than before.



Shape of power distribution

Note: A neighbor would barely notice whether two 1-MHz signals (50 watts each) had been replaced with one 2-MHz signal of 100 watts.

$$\boxed{50} + \boxed{50} = \boxed{100} \quad (\text{okay})$$

But:

$$\boxed{50} + \boxed{50} \neq \begin{array}{c} \boxed{100} \\ \boxed{100} \end{array} \quad (\text{not okay})$$

Signals cannot push power to the edges of bandwidth, creating worse adjacent-channel rejection problems for neighbors.

- Possible Rule: Average power to signal band edge cannot be more than twice the average power density of the entire signal, for all frequencies in band.



Multiple Signal Problem - 1

Problem: If each signal can radiate X unwanted signal into receiver, arbitrarily dividing a signal into two signals could allow $2X$ unwanted to be radiated.

Solution: Require all “related” signals to count as a single signal. “Related” means signals that act together, including a common signal radiated from multiple antennas, multiple carriers of a COFDM signal, etc.

Problem: If each signal can produce a ground-level signal of E_{\max} , arbitrarily dividing a signal into two signals could allow $2 \times E_{\max}$ at ground level.

Solution: Require all “related” signals to count as a single signal.



Multiple Signal Problem - 2

Problem: If each of several independent signals at a site can radiate almost E_{MAX} signal into receiver at ground level, several signals transmitting simultaneously could produce in excess of E_{MAX} signal into a receiver.

Solution: Require all signals at a single site to jointly meet the E_{MAX} limit. This requirement exists for signals regulated under the transmitter power limits or the E_{MAX} limits.

Note that this is one of the problems that also exists for sites with C&C rules today.



Relaxing the boundaries?

Problem: It may be necessary to modify the value of X in the immediate geographical vicinity of transmitters, since field strengths are very high and reducing the emission at an adjacent frequency to X would require unreasonable amounts of filtering.

Solution: Establish an emission mask for such situations that would represent a legal “safe harbor.” If the relative spectrum of the signal remained within the emission mask, it would be deemed in compliance, even if the absolute levels were above X.

Note: This is similar to many situations solved by a “site manager,” because normal rules don’t work when multiple radio systems are physically very close.



Interference?

If your system has interference, you can:

1. Show that the cause is a transmitter violating rules.

If not #1, then you can:

2. Improve your own system. Better receivers, more powerful transmitter (for better S/I).
3. Accept interference, live with it. Partial refund to clients, accept smaller area of operation, etc.
4. Negotiate with interfering transmitter for mutually acceptable business deal.

Note: Much cleaner determination of responsibility than C&C.



Modified Electrospace Rules

Only two rules:

1. Keep your signals within your licensed electrospace region (no signals greater than X outside region).
2. Restrict strong signals (limit EIRP or field strength), so that ideal receivers are not needed.

Any services, technologies, architectures, transmitter powers, modulations, etc. are permitted, as long as the two rules are obeyed. (replaces all of 47CFR and NTIA Manual)



Possible Downsides - 1

- Fewer existing rules and standards, no cookbooks.
 - More engineering expertise required?
 - Possibly more violations, due to inexperienced designs
- A more rapidly-changing signal environment
 - no expectation of protecting earlier users from more recent users. Everyone has same rules.
- Receivers must reject a wider variety of interference.
 - Receivers may be more expensive and complex compared to conventional bands, which only have a few types of interfering signals.



Possible Downsides - 2

- Who polices and documents violations of rules?
 - Possibly, spectrum users would need to hire measurement “guns” to enforce rules in cases of interference.
- Will flexible-use bands incur higher system costs
 - need to reject mix of interfering signals increases cost?
 - efficiencies from greatly increased flexibility significantly reduce costs?
- Will flexible-use typically be used to develop a new service then be abandoned for special-purpose (C&C?) bands whenever a new service gets big enough?



Flexible-Use Summary

- Flexible-use spectrum rights can provide great flexibility and rapid response to opportunities for new services and new technologies, without excessive interference among other systems.
- Responsibility for interference is well-defined.
- Market forces can provide efficient distribution of frequencies on a commodity basis via flexible secondary markets.
- Major problems seem to be completely solvable in logical ways. Some loose ends still remaining, including setting numerical values for X (leakage) and maximum power.



Final Comments

- Always a mix of regulatory environments, each needed to best support a variety of services and technologies.
- Currently unknown how the choice of numerical limits would make flexible-use bands particularly suitable (or unsuitable) for specific types of services.
- Currently unknown what eventual mix and proportion of systems in C & C, non-licensed, Part 15 underlays, and flexible-use, if all options were available.



The End

Questions
or
comments?