

Yagi Simulation: CAD-Software for Evaluation and Development

(A Case Study and Performance Report)

Rainer Bertelsmeier, DJ9BV*

Günter Hoch, DL6WU**

(Part II)

3. Yagi Basics: Function, Design, and Measurement

3.1 Radiation Mechanism

Parasitic Array Structure

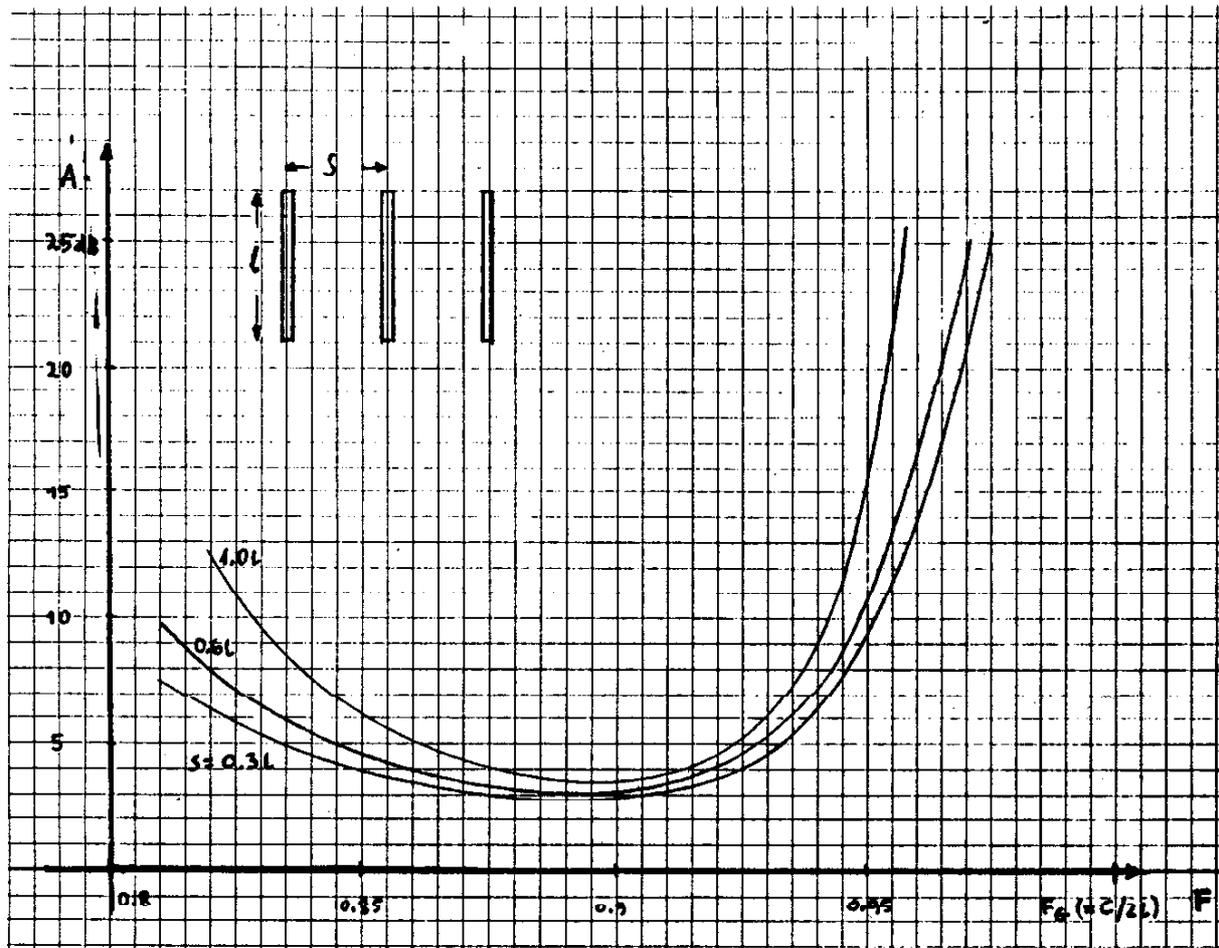
It has been shown in the previous chapter that all it takes to calculate the radiation properties of an antenna are the locations, amplitudes, and phases of all currents flowing. Thus any arbitrary arrangement of conductors excited by radiation from a driven element could be termed a parasitic array and its behavior as an antenna computed. In fact the NEC software was developed in part to do just that: calculate the influence of arbitrarily located conductors like cables, ropes, and masts on the performance of (ship) antennas. Although it is comforting to know that any conceivable structure can be analyzed this is hardly a help to the constructor in making the initial design. A closer look at a Yagi array reveals that even this seemingly simple structure has a confusing number of parameters and variables and that some knowledge of the physical principles is essential. Enough has probably been written on short Yagis and, in fact, juggling three or four elements for optimized performance is not too great a problem. As more elements are added it soon becomes evident that not all principles valid for short arrays apply to long ones as well - an approach more general than the element-by-element one is needed.

The Long Yagi Array

Looking back at the original papers by H. Yagi and S. Uda (around 1926) one finds that their principal invention was the LONG "Yagi array". They discovered that a ladder structure of dipoles slightly shorter than $1/2\lambda$, spaced less than $1/2\lambda$, acted like a waveguide and, when disrupted, like a beam antenna. They

* Gluecksburger Str. 20, D-2000 Hamburg 50

** Gersprenzweg 24, D-6100 Darmstadt



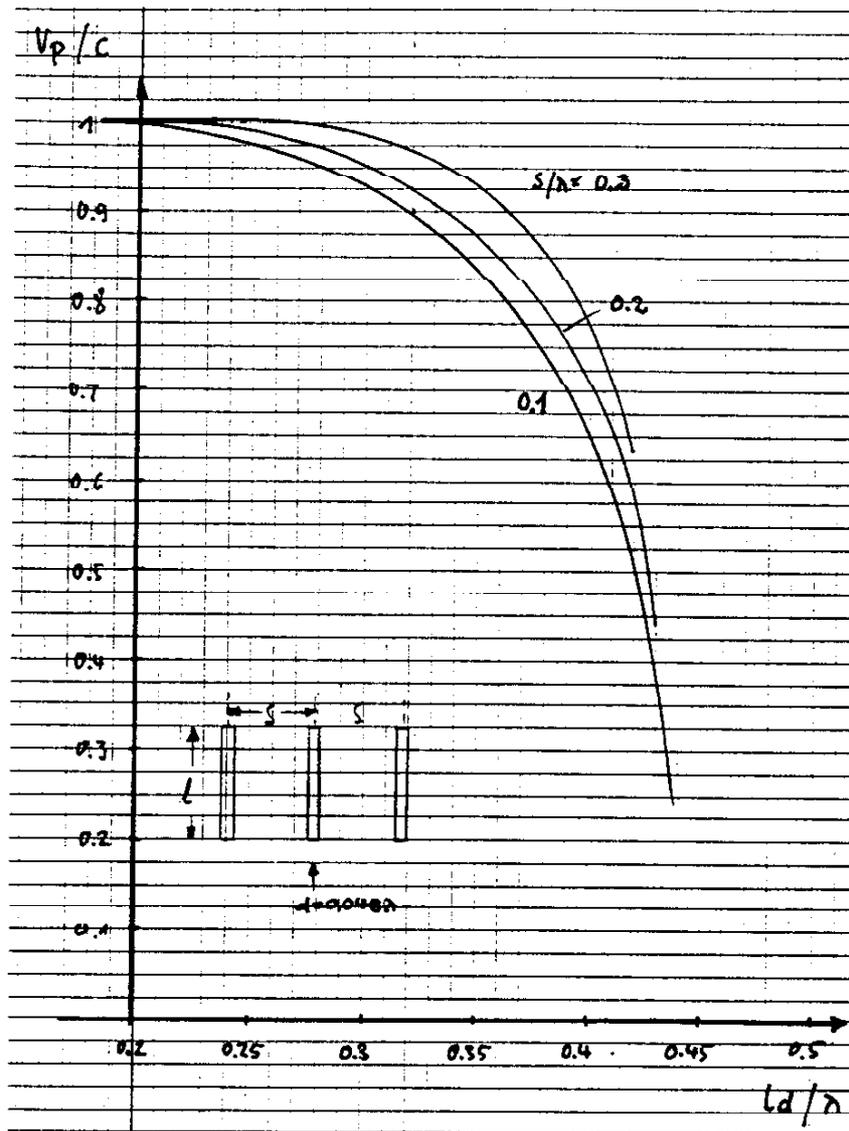
Bild/Figure 1:

Bandpass Characteristic of a typical homogeneous Yagi Waveguide versus element spacing

called this structure a "wave channel" and its elements "wave directors" (!). It is important to note that an uninterrupted Yagi structure is not an antenna. It does not radiate but conducts rf energy along the surface until it is either drained by a load or dissipated by internal loss.

Yagi Waveguides

Yagi's wave channels were homogeneous (all elements same size and spacing). Neglecting loss, on an infinite or perfectly terminated waveguide of this or any other type energy flow is the same at any point and all element currents are equal. This condition is characteristic of a travelling wave. The obvious conclusion is that to make it radiate the line must be made inhomogeneous. This can be done by either altering its parameters locally or, in the simplest case, disrupting it. For better understanding let us take a look at the geometry of a Yagi waveguide. Length and spacing of the elements determine the phase velocity of the wave propagating on it. The wave becomes slower when elements grow longer, approaching $\lambda/2$. Closer element spacing also reduces the phase velocity. The high frequency end of the passband is determined by both parameters: when elements become self-resonant ($\lambda/2$ electrically) phase velocity is



Bild/Figure 2:

Phase velocity versus elements length and spacings

reduced to zero, they turn into 'reflectors', when the spacing reaches $\lambda/2$ propagation stops through loss of coupling. The low frequency end of the pass-band is less well-defined. As elements become shorter phase velocity rises, approaching the free-space value of c , the speed of light. At the same time it becomes harder to couple energy into the line because the elements are too far from self resonance.

Artificial Dielectric

Another way to visualize the function of a Yagi structure is the analogy to a dielectric rod. The reduced phase velocity can be attributed to a dielectric constant (refractive index) greater than unity, the wave being guided by refraction into the "denser medium". Obviously this refraction ends when the dielectric constants of both media become equal, i.e. when the phase velocity on the line reaches the speed of light. There is a close analogy between long Yagis and dielectric rod antennas. The artificial dielectric model has an-

other interesting implication: characteristic impedance. The quotient of electrical and magnetical field strengths in free space is always equal to 377 Ohms, this value is called the characteristic impedance of free space. On the line the electrical field strength is reduced by a factor equivalent to the ratio of phase velocities, another way to explain why longer elements which cause a lower characteristic impedance carry higher currents. This is important in connection with efficiency because practical elements have skin effect losses which rise with the square of the current.

Discontinuous Waveguides

Armed with all this theoretical knowledge let us now return to the original Yagi waveguide. Disrupting it leads to coupling into space a fraction of the energy flow determined by the ratio of characteristic impedances, the rest being reflected and travelling back to the source. A standing wave forms on the line, radiation happens at two places: the driven element and the end of the line. This situation is characterized

by an even distribution (apart from the standing wave) of current amplitudes on the line - no energy is lost on the way. The resulting radiation pattern is rich in high sidelobes. Fig. 3 illustrates this behavior, it shows current distribution and radiation pattern of a Yagi structure with 25 identical directors.

Distributed Discontinuity

The shortcomings of the homogeneous design can be overcome by distributing the discontinuity - and thus the radiation - over the whole structure through use of a tapered length/spacing profile. Optimized designs of this type show a steady decrease of element currents from the driven element on outward. There is very little reflection from the far end because the structure performs as an exponential line transformer to the characteristic impedance of free space. Wideband low sidelobe performance results. Fig. 4 displays currents and pattern of an antenna that has the same length and number of elements as the one in fig. 3 but which is dimensioned according to the DL6WU tapering scheme. As yet there appears to be no closed theory that links optimization criteria such as pattern purity, internal loss or feed impedance with the exact distribution of currents and phases. There is, however, an old theorem by Lo, Lee and Lee that calls for a monotoneous decrease of amplitudes and phase increments over the whole length to achieve maximum gain. So, when attempting changes to a design, it is very useful to have the current distribution at hand. Experience has shown that Yagis exhibiting a "ragged" current profile usually aren't good allround performers and that elements carrying abnormally high or low currents are first choice candidates for "surgery".

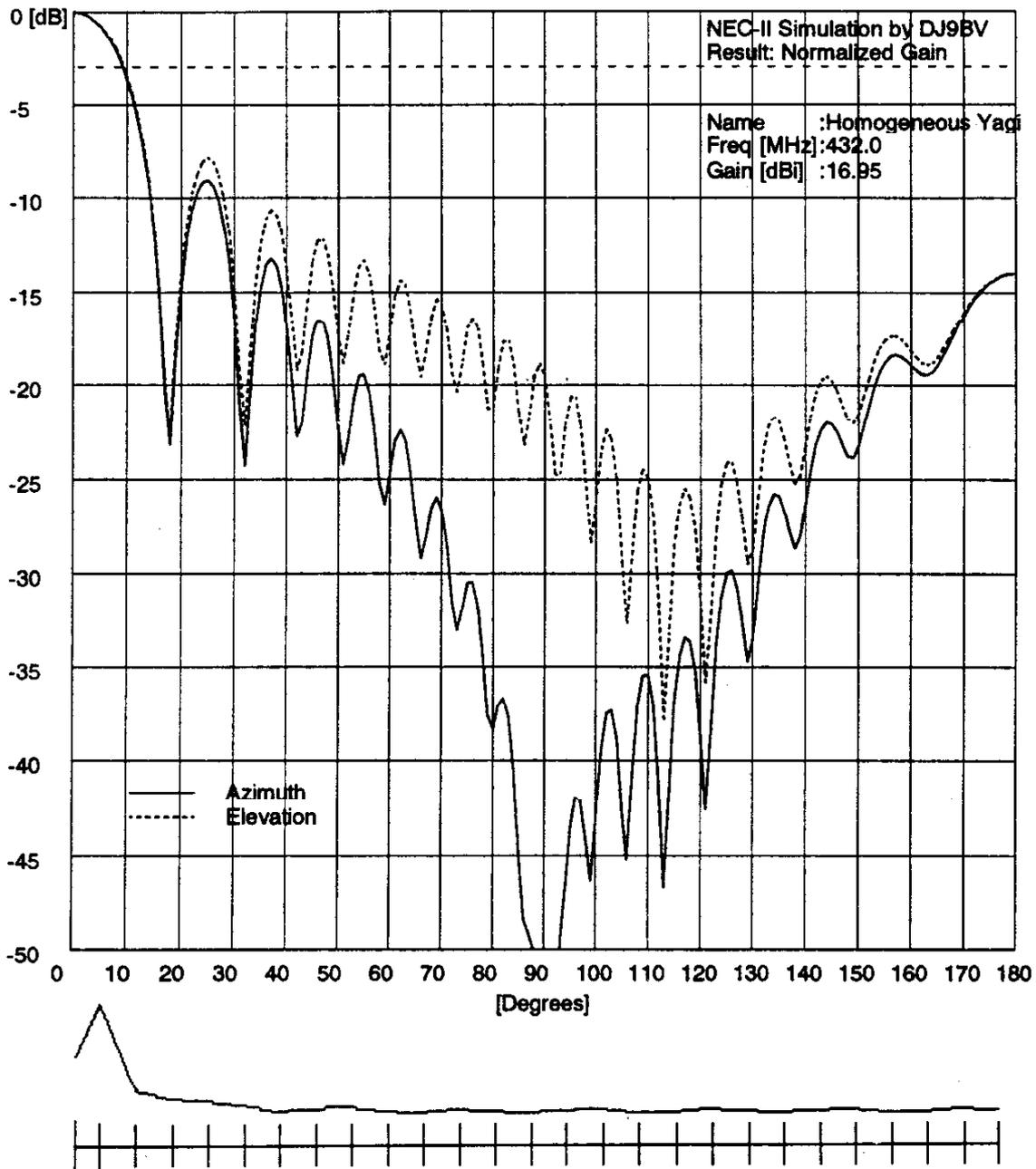
Yagi Optimization

The purpose of computer antenna modelling is almost always to find or verify an improved solution to a design problem. This means that all relevant performance data must be displayed. these include: gain (preferably with reference to some "figure of merit"), sidelobe level (not only the first but the whole panorama), F/B ratio (preferably referenced to the largest rear lobe), feed impedance (both components), efficiency (or internal loss resistance). These features are interdependent, maximum gain e.g. will not go along with extremely good values of any of the other criteria. A good modelling program should therefore inform the user of the compromises he makes when trying for enhancement of some feature. It makes little sense to compare designs unless all of the above-mentioned data are displayed. Since gain is the most important figure, means of comparing it to one of practical gain curves* should be available - there still is no theory which links maximum gain to the mechanical structure. Another point of great importance is the frequency dependence of all data, so provision must be made for altering or sweeping the operating frequency. To judge the effects of changes there should always be available data sets "before" and "after" for quick comparison. It hardly needs mention that design work is greatly alleviated by a user-friendly interface and good graphics but it should be just as clear that no serious work is possible at all if any of the essential data is not available or, perhaps even worse, only with dubious accuracy.

Limits of Optimization

It has been pointed out that no closed set of rules exist for predicting the optimum values of the diverse antenna features and their interdependence. There is, however, plenty of practical experience leading to clear conclusions.

* The gain of well designed Yagis has been determined experimentally as : $G[dBD] = 7.773*(L/wl) + 9.28$

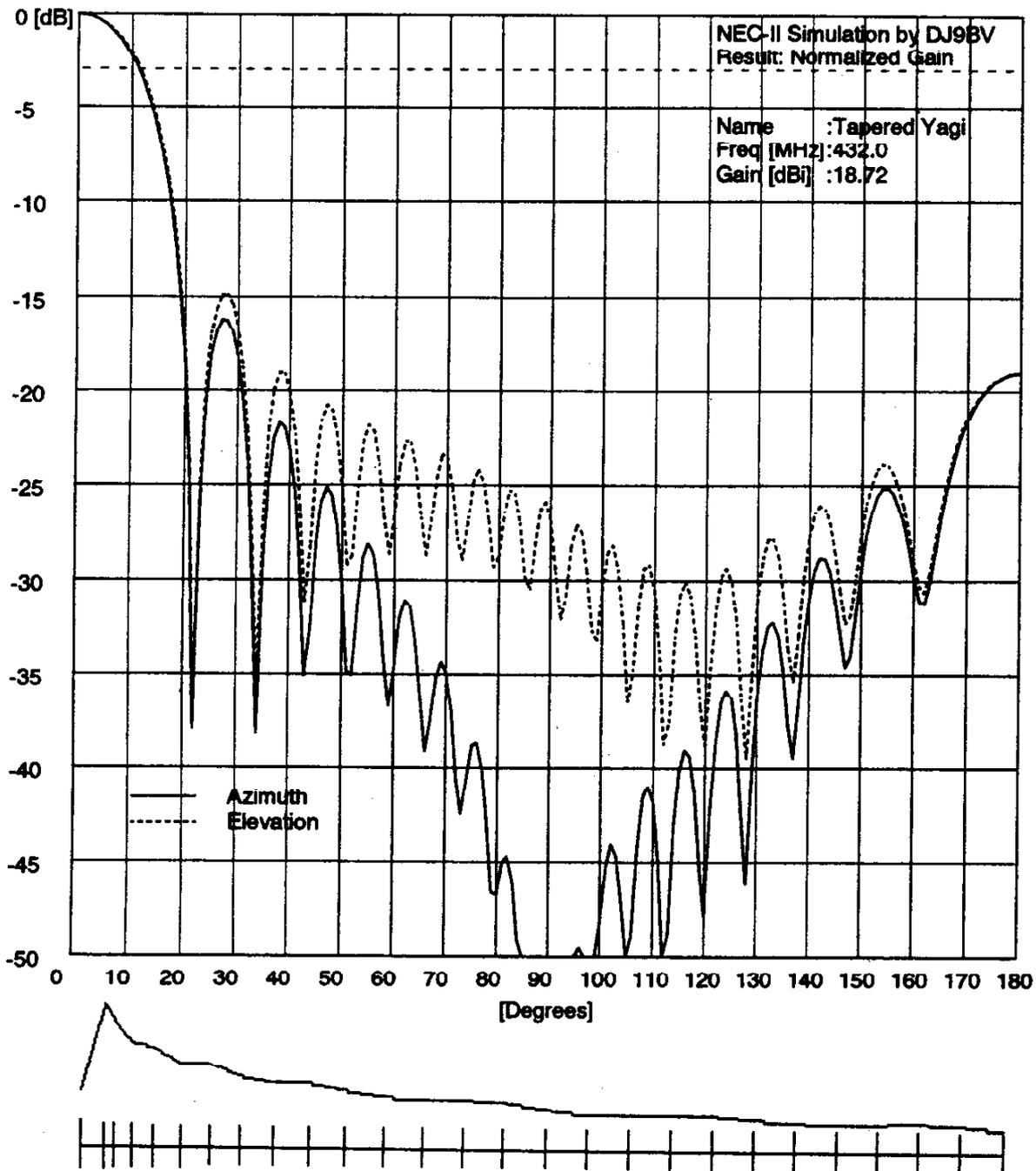


Bild/Figure 3:

Pattern and current distribution of a homogeneous Yagi

Gain

The prime goal of optimization attempts is gain - with the hope of conserving as many of the desirable secondary features as possible. Gain is related to the resultant current vector of the whole antenna as seen from the distant target. So, in some way, maximizing it always means making the element currents larger. Currents increase when the elements are spaced closer and when they approach resonance i.e. grow longer. In theory this works fine but it has adverse practical effects:



Bild/Figure 4:

Pattern and current distribution of Yagi with tapered directors (DL6WU)

lower characteristic impedance meaning lower feed impedance and higher ohmic losses. The latter are much more pronounced on VHF than on shortwaves because of the lower skin depth and the (many!) thin elements. In fact the high-frequency slope of the gain peak is almost always determined by ohmic losses starting to eat up the gain. This is why simulation programs which do not account

for skin effect give false results for height and frequency of the gain maximum. A rough estimate of the gain per length obtainable without critical side-effects is given by the DL6WU curve

Pattern

For reasons of avoiding spurious radiation and noise pickup it is desirable to reduce sidelobes as much as possible. In principle reducing sidelobe power means putting it in the main lobe but sidelobes cannot be reduced indefinitely. What is true for phased arrays applies to Yagis as well: maximum gain is obtained from an even distribution of power across the aperture and this automatically leads to a first sidelobe about 18 dB below the main lobe - analogous to the refraction rings around evenly illuminated optical apertures. All higher-order side and back lobes can (and should) be suppressed by 25 dB or better without affecting the forward gain.

Bandwidth

The maximum usable bandwidth of a Yagi antenna is always less than the passband of the underlying waveguide structure which is about 10% wide. Gain peaks somewhere within this margin leaving a practical gain bandwidth of 3...5% between -1 dB points even for the longest arrays. This is true only for tapered designs with very little end reflection. Homogeneous and other designs with high standing waves on the structure may have a much lower bandwidth. The operating bandwidth is further reduced by the matching system to the feedline. Towards the high-frequency edge of the gain curve the feed impedance drops sharply for the reasons discussed earlier. This necessitates a high transformation ratio which is inherently narrow-band and lossy. So it is often advisable to stay slightly below the gain peak to secure uncritical wide-band performance.

Figure 5 shows an example for the statements given above. Displayed are the graphs for directivity (no losses) in graph a, power gain (including losses) in graph b and actual gain (power gain - matching losses) in graph c.

Self Optimization

Some antenna design programs have a self-optimizing feature. A thorough investigation of such tools is beyond the scope of this article, only a few general remarks can be made here. All optimizers currently available are inherently "nearsighted". They all use a steepest gradient algorithm, unless they are started on the slope of the highest hill they will invariably end up on top of some minor maximum. This danger is immense with multi-element VHF/UHF Yagis although results may be quite good on few-element HF beams where the number of variables is low. These programs can be helpful in finalizing a design that is already close to optimum - another case for good handling qualities of design software because easy changes in the geometry and quick display of the consequences are necessities.

3.2 Performance Standards

To check the validity of a computer simulation or optimization it must be compared to generally accepted performance standards. This means, in effect, building the object in question and testing it on a precision antenna range. There are limits to the accuracy of such measurements - especially in the case of very long VHF/UHF Yagis which are invariably too large for anechoic chambers and must therefore be tested on open air ranges. The following sections should give an idea of the achievable precision.

Test Range Measurements

Unless special precision is required (like calibrating a gain standard) antenna measurements are made according to EIA 136. This standard imposes bounds on diverse parameters, e.g. on field homogeneity: the field must not vary by more than 0.5 dB throughout the rotation volume of the device under test. You

will hardly find an amateur test range where this basic condition is met - or even checked. From own experience we estimate the accuracy of good professional outdoor range measurements as follows:

Gain (compared to standard gain antenna)	+/- 0.2dB
Beamwidth and position of nulls os minor lobes	+/- 0.5deg
Amplitudes of side and back lobes	
0..10dB down:	+/- 0.5dB
10..20dB down:	+/- 1dB
20..30dB down:	+/- 3dB

Anything more than 30dB below the main lobe is sheer guesswork. Feed impedance can easily be measured to 1% or better.

Amateur Antenna Measurements

As pointed out in the introduction amateur measurements often differ widely from one another. There are many reasons for that. A good introduction to practical antenna measurement was given by Brown in QST ([1]), we will not repeat it here. However, the following checklist should help avoid some of the most dangerous pitfalls.

Field inhomogeneity:

caused mostly by insufficient distance from source and/or ground reflection. Results in incorrect gain and F/B readings.

Lateral field distortion:

caused by glare from reflecting objects (buildings, masts, fences etc.) inadvertently illuminated by the source. Result in pattern distortion.

Instability:

Drift of frequency, power, receiver gain, plotteroffset, ground reflectivity (rain!), etc. Results in poor repeatability of measurements and pattern plots. Can in part be overcome by ratiometric measurement.

Nonlinearity:

Deviation of amplifiers, meters, plotters etc. from specified response, mostly due to overload or insufficient SNR. Results in wrong readings and compressed patterns.

Calibration errors:

Incorrect readings due to dubious gain standards, aged or mismatched attenuators etc.

Human error:

Neglection of cable loss, mismatch, misalignment, and, and, and.

It would be unfair to state that amateur measurements approaching professional grade are downright impossible, many of the above-mentioned difficulties are pertinent to any type of antenna test work. But unless test site and equipment are checked carefully by repeated correct measurement of professionally tested antennas and gain standards the results simply can't be taken seriously.

Calibrated Computer Simulation

For the reasons just discussed it is very difficult to verify "paper" data by measurement. The advent of NEC (Numerical Electrodynamics Code) in the early eighties ([2]) raised hope of using this software as secondary standard. A copy was obtained through channels of the Research Institute of German Telecom and "tamed" by DJ9BV. The program which at that time could be run on mainframe computers only was

put through a series of tests, carefully checking the results against antenna range measurements of known precision. These were available as a byproduct of DL6WU's earlier research at the institute mentioned above. The agreement of computed and measured data turned out to be excellent. Gain and front pattern were reproduced with deviations well within the uncertainty of the open-air test range, frequency offset was less than 0.5%. Amplitudes and positions of major sidelobes were modelled to within 1dB and 0.5deg, respectively. There was a slight tendency to overestimate the first sidelobes and growing inaccuracy toward the rear pattern. Feed impedance appeared to match within 5%. Some of the deviations may have been caused by imperfections in translating physical dimensions into computer input - element holders, boom correction, dipole geometry, conductivity, etc. are difficult to model accurately. It is probably safe to say that these modelling errors are of the same order as the actual computation errors. In this light it may even be fairer to compare computation results among each other. At any rate, after critical examination, NEC can be recommended as a secondary standard. Its availability as PC software makes it an excellent and universally accessible basis of comparison, far superior to backyard or ham convention "measurements". If due care is taken in the conversion of dimensions the simulation results are comparable in quality to professional test range measurements and, another advantage, are 100% repeatable, universally. It should not be forgotten, however, that even the NEC data are not absolutely precise and that discussions about the last tenth of a dB of gain or the depth of a null are pure nonsense.

3. Yagi-Grundlagen: Funktionsweise, Entwurf und Messung

3.1 Strahlungsmechanismus

Struktur parasitärer Anordnungen

Wie im vorigen Kapitel gezeigt, genügt für die Berechnung der Strahlungseigenschaften einer Antenne die Kenntnis der Orte, Amplituden und Phasen aller darauf fließenden Ströme. Man könnte also eine beliebige Anordnung von Leitern, die von einem gespeisten Element erregt werden, als parasitäre Antenne ansehen und ihr Strahlungsverhalten ermitteln. Tatsächlich ist die NEC-Software, von der noch die Rede sein wird, zum Teil genau dafür entwickelt worden: den Einfluß beliebig angeordneter Leiter wie Kabel, Seile und Masten auf die Strahlungseigenschaften von (Schiffs-) Antennen zu berechnen. Obgleich es sehr beruhigend ist zu wissen, daß jede erdenkliche Struktur analysiert werden könnte, hilft das dem Konstrukteur herzlich wenig beim ersten Entwurf. Ein genauerer Blick auf eine Yagi-Antenne zeigt, daß diese scheinbar so einfache Struktur eine verwirrende Anzahl von Parametern und Variablen besitzt und zum Verständnis Grundkenntnisse der physikalischen Prinzipien erforderlich sind. Das gilt vor allem für lange Antennen dieses Typs; über kurze Yagis ist vermutlich genug geschrieben worden und es ist auch kein allzugroßes Problem, mit drei oder vier Elementen so lange zu jonglieren, bis ein optimiertes Verhalten erreicht ist. Wenn weitere Elemente hinzukommen, stellt sich aber bald heraus, daß nicht alles, was für kurze Antennen gilt, auch bei langen stimmt - etwas allgemeineres als die Element-für-Element-Methode wird benötigt.

Die Langyagi-Anordnung

Wenn man in die ursprünglichen Veröffentlichungen von H. Yagi und S. Uda von ca. 1926 schaut, stellt man fest, daß sie eigentlich die LANG-"Yagianordnung" erfunden haben. Sie fanden, daß eine Leiter-Anordnung von knapp $\frac{1}{2}$ langen Dipolen in weniger als $\frac{1}{2}$ Abstand sich wie ein Wellenleiter verhält und, wenn er unterbrochen wird, wie eine Richtantenne. Die nannten diese Struktur "Wellenkanal" und die Elemente "Wellendirektoren". Es ist wichtig festzuhalten, daß eine nicht unterbrochene Yagianordnung keine Antenne ist. Sie strahlt nicht, sondern leitet die HF-Energie längs ihrer Oberfläche, bis sie entweder von einer Last aufgenommen oder durch interne Verluste aufgezehrt wird.

Yagi-Wellenleiter

Yagis Wellenkanäle waren homogen (alle Elemente gleichlang und in gleichen Abständen). Läßt man die Verluste außer acht, fließt in einem solchen (oder jedem anderen) homogenen Wellenleiter an jedem Punkt die gleiche Leistung und alle Elementströme sind gleich. Dieser Zustand kennzeichnet eine Wanderwelle. Die logische Folgerung ist, daß man die Leitung inhomogen machen muß, damit sie strahlt. Das kann man durch lokale Änderung der Parameter erreichen oder im einfachsten Fall durch Abbruch der Leitung. Zum besseren Verständnis sehen wir uns noch einmal die Geometrie einer Yagi-Leitung an. Länge und Abstand der Elemente bestimmen die Phasengeschwindigkeit der sich darauf ausbreitenden Welle. Je länger die Elemente und je kleiner die Abstände werden, desto langsamer wird die Welle, ihre Geschwindigkeit wird zu null, wenn die Elemente resonant werden, also elektrisch $\lambda/2$ lang - der bekannte Effekt des "Umschlagens" von Direktoren in Reflektoren. Damit ist die obere Frequenzgrenze des Paßbandes erreicht, falls nicht schon vorher die Abstände $1/2$ Wellenlänge überschreiten, wodurch die Kopplung verloren geht. Die untere Grenze des Paßbandes ist weniger genau definiert. Wenn die Elemente im Verhältnis zur Wellenlänge kürzer werden, steigt die Phasengeschwindigkeit und nähert sich der Freiraum-Lichtgeschwindigkeit c . Gleichzeitig wird es schwieriger, Energie in die Leitung einzukoppeln und darin zu halten, da die Elemente zu weit von der Resonanz entfernt sind.

Künstliches Dielektrikum

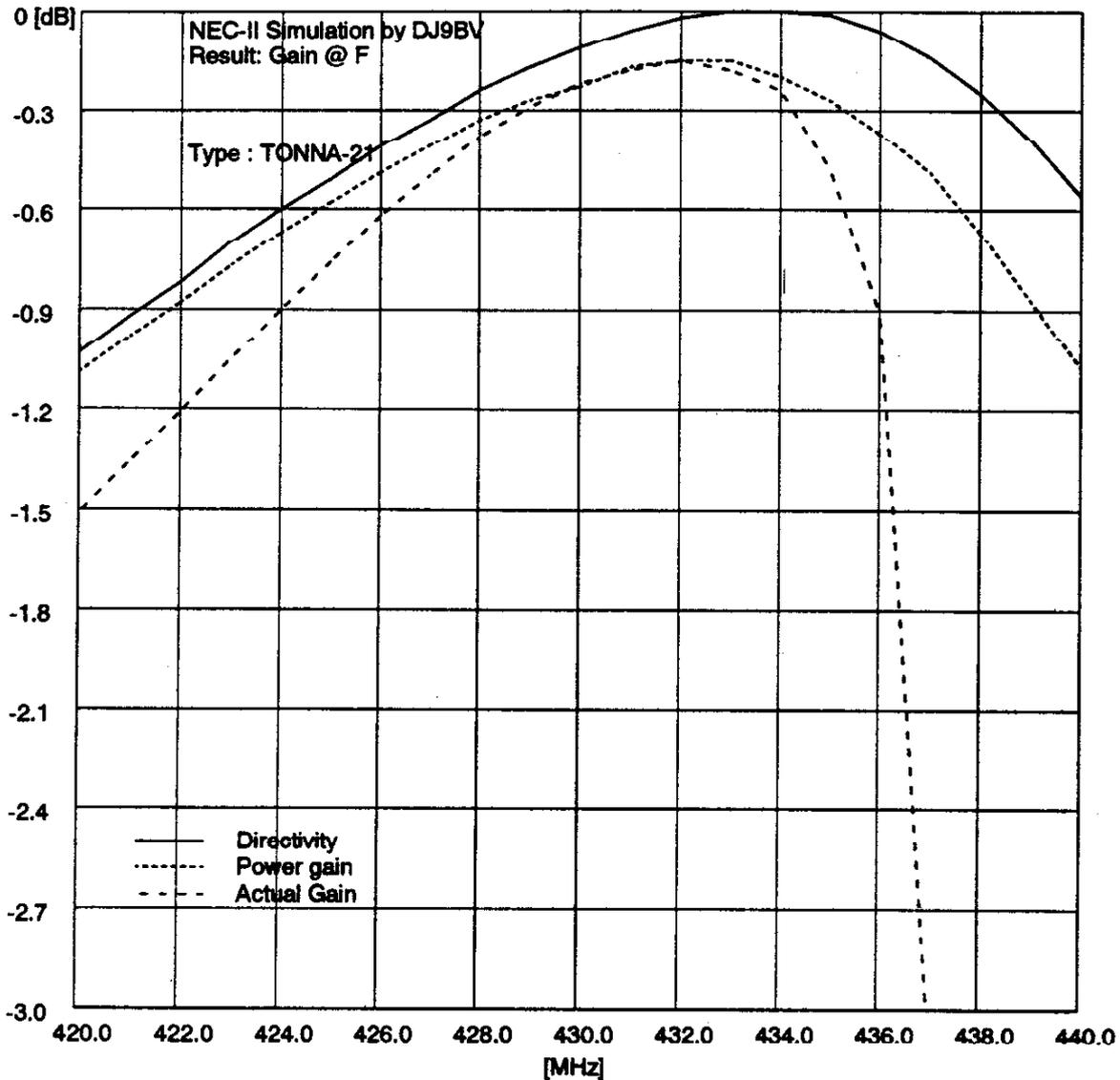
Eine andere Erklärung für die Wirkungsweise einer Yagi-Struktur beruht auf der Analogie zu einem dielektrischen Stab. Man kann die reduzierte Phasengeschwindigkeit einer Dielektrizitätskonstanten bzw. einem Brechungsindex größer als eins zuschreiben; die Welle wird geführt, indem sie zum "dichteren Medium" hin gebrochen wird. Diese Brechung hört natürlich auf, wenn die Phasengeschwindigkeit gleich der Lichtgeschwindigkeit wird, also beide Medien den gleichen Brechungsindex haben. Die enge Analogie zwischen Yagis und dielektrischen Stabantennen führt zu einer weiteren interessanten Überlegung: Im freien Raum ist der Quotient aus elektrischer und magnetischer Feldstärke stets gleich 377 Ohm , dem Freiraumwellenwiderstand. Auf der "Leitung" ist er infolge der höheren Dielektrizitätskonstanten reduziert und zwar um einen Faktor, der dem Verhältnis der Phasengeschwindigkeiten entspricht. Diese Betrachtungsweise erklärt auch zwanglos, warum bei elektrisch länger werdenden Elementen, also niedrigerem Wellenwiderstand, die Ströme anwachsen. Dieser Punkt ist im Zusammenhang mit den inneren Verlusten von Bedeutung, da ja auf praktischen Leitern Skineffektverluste entstehen, die mit dem Quadrat des Stroms zunehmen.

Unterbrochene Wellenleiter

Mit einer ganzen Menge Theorie bewaffnet kehren wir nun zu den ursprünglichen Yagi-Wellenleitern zurück. Es wird jetzt verständlich, warum an einer Abbruchstelle ein Teil des Energieflusses, der durch das Verhältnis der Wellenwiderstände bestimmt wird, in den Raum abgegeben wird, während der Rest reflektiert wird und zum Erreger zurück fließt. Auf der Leitung bildet sich eine stehende Welle aus, Abstrahlung passiert nur an den zwei Diskontinuitäten: am Strahler und am offenen Leitungsende. Auf der Leitung herrscht abgesehen von der stehenden Welle eine gleichmäßige Verteilung der Stromamplituden, es geht auf dem Weg (außer durch ohmsche Verluste) keine Energie verloren. Das resultierende Strahlungsdiagramm ist reich an großen Nebenzipfeln. Bild 3 illustriert dieses Verhalten, es zeigt Stromverteilung und Strahlungsdiagramm einer homogenen Yagiantenne mit 25 Direktoren.

Verteilte Diskontinuität

Die Nachteile der homogenen Struktur lassen sich vermeiden, indem man die Diskontinuität und damit die Abstrahlung über die ganze Länge verteilt. Man erreicht dies durch ein kontinuierlich gestuftes Profil der Elementlängen und Abstände. Auf einer optimierten Struktur dieser Art nehmen die Elementströme



Bild/Figure 5:

TONNA-21 (Yagi for 432 MHz): Gain versus Frequency

Graph a: Directivity D

Graph b: Power Gain $P = D \times \eta$ (η Efficiency)

Graph c: Actual gain = $P \times ML$ ($ML = 1 - r^2$)

vom Strahler an zum Antennenende hin stetig ab. Es tritt am offenen Ende kaum Reflexion auf, da die Anordnung als Exponentialtransformator zum Wellenwiderstand des freien Raums wirkt. Bandbreite und Nebenzipfelunterdrückung wachsen immens. Bild 4 zeigt Ströme und Diagramm einer Antenne, die gleiche Länge und Elementanzahl aufweist, wie die in Bild 3, aber nach dem DL6WU-Profil abgestuft ist. Bis jetzt gibt es keine vollständige Theorie für die Zusammenhänge zwischen den Optimierungskriterien wie Richtdiagramm, Wirkungsgrad und Speisewiderstand und der exakten Verteilung der Ströme und Phasen. Es existiert allerdings eine alte Theorie von Lo, Lee und Lee, die für maximalen Gewinn von Längsstrahlern eine monotone Abnahme der Amplituden und Phasendifferenzen über die ganze Länge

fordert. Deshalb ist es wichtig, die Stromverteilung angezeigt zu sehen, wenn man Veränderungen an einem Entwurf vornimmt. Die Erfahrung hat gezeigt, daß Yagis mit einem zerklüfteten Stromprofil selten gute Allroundeigenschaften besitzen und daß die Elemente mit abnorm hohen oder niedrigen Strömen die ersten Kandidaten für "chirurgische Eingriffe" sind.

Optimieren von Yagi-Antennen

Modellrechnungen an Antennen dienen fast immer dem Zweck, eine verbesserte Lösung für ein Entwurfsproblem zu finden oder zu verifizieren. Das bedeutet, daß alle relevanten Daten angezeigt werden müssen, nämlich Gewinn, Nebenzipfeldämpfung (nicht nur der erste, sondern das ganze Panorama), Vor-Rück-Verhältnis (am besten in Bezug auf den größten Rückzipfel), Speiseimpedanz (beide Komponenten) und Wirkungsgrad bzw. interner Verlust. Diese Zahlen sind nicht unabhängig voneinander, maximaler Gewinn ist beispielsweise nicht verträglich mit extrem guten Werten der anderen Daten. Ein gutes Simulationsprogramm sollte deshalb den Benutzer über die Kompromisse informieren, die er für die Verbesserung einer Größe in Kauf nimmt. Es ist auch nicht sinnvoll, verschiedene Entwürfe zu vergleichen, wenn nicht alle genannten Werte zur Verfügung stehen. Der Gewinn als wichtigste Optimierungsgröße sollte zusätzlich im Vergleich mit einer der gängigen Gewinn-Längen-Kurven angezeigt werden, da es noch keine gültige Formel für den erzielbaren Maximalgewinn gibt. Ein weiterer wichtiger Punkt ist die Frequenzabhängigkeit aller Daten, deshalb muß die Möglichkeit bestehen, die Arbeitsfrequenz zu verschieben oder zu wobbeln. Um die Auswirkung von Änderungen zu beurteilen sollten die Daten "vorher" und "nachher" leicht vergleichbar zur Verfügung stehen. Es bedarf kaum der Erwähnung, daß eine benutzerfreundliche Bedieneroberfläche und gute Graphik die Arbeit enorm erleichtern, aber es sollte ebenso klar sein, daß eine sinnvolle Optimierungsarbeit überhaupt nicht möglich ist, wenn auch nur einer der wichtigen Werte fehlt oder - vielleicht noch schlimmer - mit zweifelhafter Genauigkeit ausgegeben wird.

Grenzen der Optimierung

Wie schon gesagt gibt es keine endgültigen Regeln für die Maximalwerte der verschiedenen Kenngrößen und ihre gegenseitige Abhängigkeit, es gibt jedoch so viel praktische Erfahrung, daß konkrete Aussagen möglich sind.

Gewinn:

Das Hauptziel jeder Optimierung ist der Gewinn, immer in der Hoffnung, dabei möglichst viele der sekundären Ziele zu erhalten. Der Gewinn wird durch den resultierenden Stromvektor der ganzen Antenne aus der Sicht des entfernten Ziels bestimmt; ihn zu maximieren bedeutet immer in irgend einer Weise, die Elementströme zu vergrößern. Die Ströme erhöhen sich mit Annäherung an die Resonanz, also Verlängerung, und durch geringere Abstände. Das funktioniert in der Theorie bestens, hat aber in der Praxis unangenehme Folgen: niedrigeren Wellenwiderstand und damit auch Speisewiderstand und höhere ohmsche Verluste. Besonders diese machen sich im UKW-Bereich weit mehr bemerkbar als auf Kurzwellen, weil die Eindringtiefe kleiner ist und die (vielen!) Elemente dünner sind. Tatsächlich wird die gemessene Gewinn-Frequenzkurve nahezu immer durch die Verluste eingeengt, die auf der hochfrequenten Flanke anfangen, den Gewinn aufzuzehren. Aus diesem Grund liefern alle Programme, die den Skineneffekt nicht berücksichtigen, falsche Werte für Frequenz und Höhe des Gewinnmaximums. Eine grobe Abschätzung des ohne kritische Nebeneffekte erzielbaren Gewinns liefert die DL6WU-Gewinnkurve (s.u.).

* Der experimentell ermittelte Gewinn von DL6WU-Yagis folgt der Gleichung: $G [dBD] = 7,773 \cdot \text{LOG}(L/WL) + 9,28$

Richtdiagramm:

Zur Vermeidung von Störstrahlung und -empfang ist es wünschenswert, die Nebenkeulen so weit wie möglich zu unterdrücken. Im Prinzip heißt Reduzierung der Nebenzipfel mehr Energie in der Hauptkeule, aber das hat Grenzen. Bei Flächen- und Aperturantennen führt eine gleichmäßige Verteilung der Strahlungsleistung über die ganze Fläche zu maximalem Gewinn und automatisch zu einem ersten Nebenzipfel von -18 dB - analog zu den unvermeidlichen Beugungsringen in der Optik. Ähnliches gilt auch für Yagis. Wenn der erste Nebenzipfel weniger als 15 dB oder mehr als 20 dB unter der Hauptkeule liegt, ist die Antenne mit Sicherheit nicht optimal. Alle Nebenkeulen höherer Ordnung sowie die Rückkeulen können (und sollten) auf -25 dB oder mehr abgesenkt werden, ohne den Gewinn zu beeinträchtigen.

Bandbreite:

Die maximal nutzbare Bandbreite einer Yagiantenne ist immer kleiner als das Durchlaßband des zugrundeliegenden Yagi-Wellenleiters, d.h. etwa 10%. Innerhalb dieses Bandes erreicht der Gewinn sein Maximum, zwischen den -1 dB-Punkten verbleiben auch bei den längsten Antennen 3-5%, was allerdings nur für gestufte Profile mit geringer Endreflexion gilt. Die Arbeitsbandbreite wird weiter eingeschränkt durch die Anpassung an die Speiseleitung. Auf der hochfrequenten Seite der Gewinnkurve fällt der Speisewiderstand, wie schon erläutert, schnell ab. Das erfordert ein höheres Transformationsverhältnis, der Frequenzbereich wird kleiner und die Verluste größer. Trotz des geringfügig niedrigeren Gewinns ist es deshalb meist ratsam, die Arbeitsfrequenz etwas unterhalb des Gewinnmaximums zu legen. Es ist übrigens ein Irrglaube, daß der Elementdurchmesser an all diesen Effekten nennenswert etwas ändert. Auch der Gewinn ist davon weitestgehend unabhängig.

Wirkungsgrad und Fußpunktimpedanz:

Versuche, das letzte Zehntel-dB aus einer Yagi herauszuquetschen, führen unweigerlich zu hohen Elementströmen und niedriger Speiseimpedanz. Daraus resultieren erhöhte Skineneffektverluste, die nicht nur den Gewinn mindern, sondern bei Low-Noise-Anwendungen wie EME zu zusätzlichem Rauschen führen. 0.2 dB mehr ohmsche Verluste können leicht 1 dB geringere Empfindlichkeit des Systems verursachen! Die unerwünschten Effekte der niedrigen Speiseimpedanz wurden bereits erwähnt. Als praktische Regel wird empfohlen, die ohmschen Verluste auf 0.1 dB zu begrenzen und den Fußpunktwiderstand (an einem offenen Dipol) nicht unter 25 Ohm sinken zu lassen. "Optimierungen", die Speise- und Verlustwiderstand außer acht lassen, sind Zeitverschwendung.

Die o.a. Zusammenhänge werden beim Betrachten der Gewinnbandbreite einer TONNA-21 in Bild 5 belegt. Kurve a zeigt den Verlauf der Richtwirkung (ohne Verluste!) mit der Frequenz, Kurve b zeigt den Leistungsgewinn (berücksichtigt die Verluste) und Kurve c zeigt schließlich den Leistungsgewinn abzüglich Anpaßverlust. Diese Kurven wurden durch Simulation mit NEC-II gewonnen.

Automatische Optimierung

Einige Antennenberechnungsprogramme bieten eine Funktion "Automatische Optimierung" an. Eine gründliche Analyse dieser Werkzeuge sprengt den Rahmen dieser Abhandlung, hier können nur ein paar allgemeine Bemerkungen dazu gemacht werden. Alle derzeit verfügbaren Optimierer sind "kurzsichtig". Sie arbeiten nach einem Gradientensuchverfahren; wenn sie nicht bereits auf der Flanke des höchsten "Gewinnberges" gestartet werden, landen sie unweigerlich auf einem Nebenmaximum. Diese Gefahr ist bei vielelementigen UKW-Yagis extrem groß, auch wenn das gleiche Programm bei KW-Antennen mit wenigen Elementen (und Variablen) vernünftige Resultate liefern mag. Optimierer können allerdings beim Feinschliff eines schon fast ausgereizten Entwurfs sehr hilfreich sein. Hier wird wieder deutlich, wie wichtig eine gute Benutzeroberfläche für eine Entwurfssoftware ist, weil manuelle Eingriffe und schnelle Anzeige der Auswirkungen erforderlich sind.

3.2 Leistungsvergleiche

Wenn man die Güte einer Simulation oder Optimierung überprüfen will, muß man sie mit allgemein akzeptierten Bezugsdaten vergleichen. Im Prinzip müßte man also die betreffende Antenne aufbauen und auf einer Präzisionsmeßstrecke vermessen. Auch solche Messungen haben allerdings ihre Grenzen. Beispielsweise sind lange VHF/UHF-Yagis in aller Regel zu groß für Absorberhallen und müssen deshalb im Freifeld gemessen werden. Die nächsten Abschnitte sollen Anhaltspunkte für die zu erwartende Genauigkeit geben.

Freifeldmessungen

Wenn nicht besondere Anforderungen an die Genauigkeit gestellt werden, wie bei der Kalibrierung von Gewinnnormalen, wird auf Antennenmeßplätzen nach der EIA-Norm 136 verfahren. Diese Norm legt Mindestanforderungen an diverse Parameter fest, so beispielsweise an die Feldhomogenität: weniger als 0.5 dB Variation im Drehvolumen des Meßobjekts. Man findet kaum eine Amateur-Meßstrecke, bei der diese Grundforderung erfüllt oder auch nur geprüft wird. Aus eigener Erfahrung schätzen wir die Genauigkeit, die bei guten professionellen Freifeldmessungen erzielt wird, etwa wie folgt:

Gewinn (verglichen mit Gewinnnormal)	+/-0.2 dB
Öffnungswinkel und Lage von Nebenkeulen	+/-0.5 Grad
Amplituden von Neben- und Rückkeulen im Bereich	
0...-10 dB	+/- 0.5 dB
-10...-20 dB	+/- 1 dB
-20...-30 dB	+/- 3 dB

Details, die mehr als 30 dB unter der Hauptkeule liegen, können nur als Schätzungen gelten. Der Fußpunktwiderstand kann leicht auf 1% genau gemessen werden.

Amateurmessungen

Wie bereits in der Einleitung erwähnt, weichen Amateur-Antennenmessungen oft weit voneinander ab. Dafür gibt es viele Gründe. Eine gute Einführung in die Antennenmeßpraxis hat F. Brown in der QST ([1]) veröffentlicht, sie soll hier nicht wiederholt werden. Die folgende Checkliste sollte aber helfen, die größten Fehler zu vermeiden.

Feldinhomogenität:

Wird meist durch ungenügenden Abstand zur Signalquelle und/oder Bodenreflexionen hervorgerufen. Verursacht ungenaue Gewinn- und F/B-Werte.

Seitliche Feldverzerrung:

Wird durch Reflexionen an spiegelnden Objekten wie Gebäude, Zäune, Masten etc. erzeugt, die ungewollt von der Signalquelle mitangestrahlt werden. Ergibt verzerrte Diagramme.

Instabilität:

Drift von Frequenz, Signalleistung, Empfänger-Verstärkung, Schreiber-Nullage, Bodenreflektivität (Feuchtigkeit!) etc. zeigt sich in schlechter Reproduzierbarkeit von Meßwerten und Diagrammen. Kann zum Teil durch ratiometrische Meßmethoden ausgeschaltet werden.

Nichtlinearität:

Abweichung von Verstärkern, Meßinstrumenten, Schreibern usw. von der angegebenen Kennlinie, meist durch Übersteuerung oder ungenügendes Signal-Rausch-Verhältnis bedingt. Ruft falsche Meßwerte und gestauchte Diagramme hervor.

Eichfehler:

Fehlmessungen durch zweifelhafte Gewinnnormale, alte oder fehlangepaßte Dämpfungsglieder o.ä.

Menschliche Fehler:

Nichtberücksichtigung von Fehlanpassung, Kabeldämpfung, Ausrichtfehler und, und, und.

Es wäre unfair, zu behaupten, daß Amateurmessungen mit fast professioneller Genauigkeit schlichtweg unmöglich sind, viele der hier geschilderten Fehlerquellen treten bei jeder Art von Messung auf. Aber wenn nicht die Meßstrecke und die Geräte durch sorgfältige Testmessungen, am besten mittels professionell vermessener Antennen und Gewinnnormale, mehrfach überprüft und für korrekt befunden wurden, kann man die Ergebnisse nicht ernst nehmen.

Kalibrierte Computersimulation

Aus den genannten Gründen ist es schwierig, "Papierdaten" durch Messungen zu verifizieren. Das Erscheinen der NEC-Software (Numerical Electromagnetics Code) ([2]) anfangs der achtziger Jahre weckte Hoffnungen, dieses Programm als sekundären Maßstab heranziehen zu können. Durch Kanäle des Forschungsinstituts der Deutschen Bundespost konnte eine Kopie dieses erstaunlicherweise nicht geheimen Werkes beschafft und durch DJ9BV "gezähmt" werden. Das Programmpaket, das seinerzeit nur auf Großrechnern lauffähig war, wurde dann einer Reihe sorgfältiger Tests unterzogen. Dabei wurden die Rechenergebnisse mit Meßplatzresultaten bekannter Genauigkeit verglichen, die als Abfallprodukt der früheren Arbeit DL6WU's am genannten Institut vorlagen. Die Übereinstimmung erwies sich als hervorragend. Der Gewinn und die Hauptkeule wurden mit einer Genauigkeit reproduziert, die innerhalb der Meßgenauigkeit eines guten Freifeldmeßplatzes lag, Frequenzabweichungen waren geringer als 0.5%. Die Amplituden und Positionen der größeren Nebenkeulen stimmten auf 1 dB bzw. 0.5 Grad mit den Messungen überein. Dabei war eine leichte Tendenz zur Überschätzung der ersten Nebenkeule und eine deutliche Abnahme der Übereinstimmung mit steigender Winkelablage von der Hauptkeule zu beobachten. Die Speiseimpedanz zeigte Übereinstimmung auf ca. 5%. Einige der Abweichungen sind sicher damit zu erklären, daß bereits die Übersetzung der mechanischen Dimensionen einer Antenne in Computer-Input Fehlerquellen enthält - so sind beispielsweise Elementhalter, Boomkorrektur, Dipolgeometrie, Leitfähigkeit nur schwer im Modell zu erfassen. Man muß vermuten, daß diese Übersetzungsfehler bereits in der gleichen Größenordnung wie die internen Rechenungenauigkeiten liegen. In diesem Licht besehen wäre es möglicherweise sogar fairer, nur Computerergebnisse untereinander zu vergleichen. Auf jeden Fall kann nach dieser kritischen Untersuchung NEC als sekundärer Vergleichsmaßstab empfohlen werden. Da es inzwischen auch als PC-Software erhältlich ist, stellt damit eine weltweit einheitliche Vergleichsbasis zur Verfügung, die jeder Hinterhof- oder Hamtagungs-"Messung" haushoch überlegen ist. Bei sorgfältiger Konversion der Eingabedaten sind die Ergebnisse von einer Qualität, die mit professionellen Messungen vergleichbar ist und, ein weiterer Vorteil, weltweit jederzeit 100% reproduzierbar sind. Trotzdem sollte nie vergessen werden, daß auch die NEC-Berechnungen nur endlich genau sind und deshalb Streiterei um das letzte Zehntel-dB Gewinn oder die Tiefe einer Nullstelle unsinnig ist.

References/Literatur

[1] F. Brown, "Antenna Gain Measurements", QST, November 1982, pp. 35-37 and December 1982, pp.27-31

[2] G.J. Burke, A.J. Poggio, "Numerical Electromagnetics Code (NEC) - Method of Moments", NOSC TD 116 , Vol. 1: Program Description - Theory, Vol. 2: Program Description - Code, Vol. 3: Users's Guide, San Diego, January 1981