How to Tangle with a Nested Radical

Susan Landau*

Introduction

Like many an intriguing question in algebraic manipulation, the problem of denesting nested radicals had its origins with Ramanujan. That is not to say that no one had ever considered the problem of denesting radicals before he did. Certainly, the fact that

$$\sqrt{5+2\sqrt{6}} = \sqrt{2}+\sqrt{3}$$

is simple enough that it must have been known several centuries ago. Ramanujan [11] upped the ante. For each of the formulas below, he took the doubly nested radical on the left and simplified it to a combination of singly nested radicals on the right:

$$\begin{aligned}
\sqrt[3]{\sqrt[3]{2}-1} &= \sqrt[3]{1/9}-\sqrt[3]{2/9}+\sqrt[3]{4/9},\\
\sqrt{\sqrt[3]{5}-\sqrt[3]{4}} &= (1/3)(\sqrt[3]{2}+\sqrt[3]{20}-\sqrt[3]{25}),\\
\sqrt[6]{7\sqrt[3]{20}-19} &= \sqrt[3]{5/3}-\sqrt[3]{2/3},\\
\sqrt[4]{\frac{3+2\sqrt[4]{5}}{3-2\sqrt[4]{5}}} &= \frac{\sqrt[4]{5}+1}{\sqrt[4]{5}-1},\\
\sqrt{\sqrt[3]{28}-\sqrt[3]{27}} &= (1/3)(\sqrt[3]{98}-\sqrt[3]{28}-1),\\
\sqrt[3]{\sqrt[5]{32/5}-\sqrt[5]{27/5}} &= \sqrt[5]{1/25}+\sqrt[5]{3/25}-\sqrt[5]{9/25}.
\end{aligned}$$

What Ramanujan neglected to do was provide a theory for simplifying nested radicals. When computers came along, symbolic computation became important. There was a practical reason to find an algorithm for denesting nested radicals.

A machine has no problem with

$$1,\ \sqrt{5+2\sqrt{6}},\ 5+2\sqrt{6},\ (5+2\sqrt{6})^{3/2}$$

as a basis for $Q(\sqrt{5+2\sqrt{6}})$ over Q. Most human beings seem to prefer the basis

$$1,\ \sqrt{2},\ \sqrt{3},\ \sqrt{6}.$$

The difficulty is that there was no general method to go from the complex form of a nested radical to a simplified version. If Ramanujan had one, he never wrote it down.

Necessity has often been the mother of invention, and so it proved to be in this case. Although the general problem remains open, there are now solutions to a number of subproblems: for denesting real nested square roots [3], for denesting real radicals of depth 2 [5], for radicals of a special form [9, 13], and for radicals using roots of unity [7, 8]. I am interested in three questions: When does a simplification exist? Is there a technique for finding it? How long does it take? In this article, I will briefly present some recent theorems for radical simplification, and the algorithms they lead to. For proofs, and complete presentations, the reader is urged to read the original papers.

What Does It Mean to Denest a Radical?[1]

I begin by making precise what is meant by simplifying a nested radical. Assume all fields are characteristic 0. In defining the *depth of nesting* of a formula, I will view the formula as a sequence of formal symbols. Following [3], a *formula* over a field k and its *depth* of nesting are defined recursively:

- An element in the field k is a formula of depth 0 over k. *Thus,* 17 *is of depth* 0 *over* Q, *while* $1+\sqrt{2}$ *is of depth* 0 *over* $Q(\sqrt{2})$.

Susan Landau

Susan Landau received her Ph.D. from MIT in 1983. Since then she has taught at Wesleyan University and the University of Massachusetts at Amherst. Her research interests include computational number theory, algebraic algorithms, and symbolic computation. When she is not proving theorems or reading e-mail, she likes to attend the theater or the New York City Ballet, or to go hiking with her dog, husband, and two children.

[1] This brief overview is taken from [8]. A more detailed discussion of these issues can be found in [7].

* Supported by NSF grant DMS-8807202.

- An arithmetic combination (A ± B, A × B, A/B) of formulas A and B is a formula whose depth over k is max(depth(A), depth(B)). *Because I view $\sqrt{5+2\sqrt{6}} - \sqrt{2} - \sqrt{3}$ as a sequence of formal symbols, it has nesting depth* 2 *over Q.*
- A root $\sqrt[n]{\mathrm{A}}$ of a formula A is a formula whose depth over k is 1+ depth(A).
- *Finally, $\sqrt{\sqrt{5+2\sqrt{6}}}$ has depth* 3 *over Q.*

Such a formula is called a nested radical. A nesting of α means any formula A that can take α as a value. But there are difficulties involved as an nth root is a multivalued function. When I write the equation

$$\sqrt{5+2\sqrt{6}} = \sqrt{2} + \sqrt{3},$$

it is unclear which $\sqrt{2}$ I mean and which $\sqrt{3}$. The usual interpretation is the positive real roots for all four choices in the equation above. Under those choices, the equation is correct; under others, it may not be.

I start with the input as a sequence of expressions of the form:

$$\begin{aligned}
\alpha_1 &= \sqrt[n_1]{q},\ q \in k,\\
\alpha_2 &= \sqrt[n_2]{\tilde{p}_2(\alpha_1)},\ \tilde{p}_2 \in k[x_1],\\
\alpha_3 &= \sqrt[n_3]{\tilde{p}_3(\alpha_1,\alpha_2)},\ \tilde{p}_3 \in k[x_1,x_2],\\
&\ \ \vdots\\
\alpha_m &= \tilde{q}(\alpha_1,\ldots,\alpha_{m-1}) + \sqrt[n_m]{\tilde{p}_m(\alpha_1,\ldots,\alpha_{m-1})},\\
&\tilde{q} \text{ and } \tilde{p}_m \in k[x_1,\ldots,x_{m-1}] \text{ and } \alpha = \alpha_m.
\end{aligned}$$

It is not hard to go from this complicated sequence to the minimal polynomial for α over k. One can do it by first determining a minimal polynomial for α_1 over k, then using that to determine a minimal polynomial for α_2 over k, and so on (see [8] for details). One must take careful note of the choices of roots as they are made.

Once one chooses a particular nth root for $\sqrt[n]{\alpha}$, the same value must be assigned to it each time it appears. If the roots are specified at the time a nested radical is given, choose those roots. Whenever roots appear which have not been previously specified, one is free to pick a value arbitrarily for them, so long as after that one consistently chooses the same value to represent the root.

Suppose one is interested in denesting the expression

$$\sqrt[3]{\sqrt[3]{2}-1} - \sqrt[3]{1/9}.$$

The polynomial x^3-9 factors over the field $Q(\sqrt[3]{\sqrt[3]{2}-1})$. To denest $\sqrt[3]{\sqrt[3]{2}-1} - \sqrt[3]{1/9}$, I need to know to which root of x^3-9 I am referring in $Q(\sqrt[3]{\sqrt[3]{2}-1})$: the one which satisfies $x - \alpha^8 - 4\alpha^5 - 4\alpha^2$, or one of the two satisfying $x^2 + (\alpha^8+4\alpha^5+4\alpha^2)x + (3\alpha^4+6\alpha)$, where $\alpha = \sqrt[3]{\sqrt[3]{2}-1}$.

For the purposes of this paper, I have chosen that when I adjoin $\sqrt[n]{\alpha}$, I do so in a way that makes the smallest (in terms of degree) field extension possible. In the above example, I would choose the $\sqrt[3]{9}$ that is already in the field.

I will say the formula A *can be denested over the field k* if there is a formula B of lower depth than A such that A and B have the same (real or complex) value. I will say that A *can be denested in the field L* if there is a formula B of lower nesting depth than A with all of the terms (subexpressions) of B lying in L, again with A and B having the same value. For any α, I define the depth of α over k to be the depth of the minimum-depth expression for α. When I am given a formula A for α such that A can be denested, I will sometimes say that α can be denested.

For the remainder of this article I will assume that α has been given by its minimal polynomial over k, and the choice of roots in any ambiguous situation has been spelled out.

Denesting Real Nested Radicals

Real nested radicals were Ramanujan's examples and form a clear starting point for the problem. Nested square roots form the simplest example of nested radicals. Since nested real square roots describe the Euclidean distance from one vertex on a polyhedron to another, an algorithm for their simplification is potentially of practical value. With such concerns in mind, Borodin, Fagin, Hopcroft, and Tompa [3] studied simplifying nested square roots. Their first result demonstrates when square roots suffice for denesting:

THEOREM 1 [3]. *Let $Q \subseteq k$, and let a, b, r be elements of k, with $\sqrt{r}$ not in k. The the following are equivalent:*

1. *$\sqrt{a+b\sqrt{r}}$ is in $k(\sqrt{r}, \sqrt{a_1}, \ldots, \sqrt{a_l})$ for some $a_1, \ldots, a_l$ in k.*
2. *$\sqrt{a+b\sqrt{r}}$ is in $k(\sqrt{s}, \sqrt{r})$ for some $s \neq 0$ in k.*
3. *$\sqrt{a^2-b^2r}$ is in k.*

Next they gave the conditions under which fourth roots may help:

THEOREM 2 [3]. *Let $Q \subseteq k$, and let a, b, r be in k, with $\sqrt{r}$ not in k. Then the following are equivalent:*

1. *$\sqrt{a+b\sqrt{r}}$ is in $k(\sqrt[4]{r}, \sqrt{a_1}, \ldots, \sqrt{a_l})$ for some $a_1, \ldots, a_l$ in k.*
2. *Either $\sqrt[4]{r}\sqrt{s}\sqrt{a+b\sqrt{r}}$ or $\sqrt{s}\sqrt{a+b\sqrt{r}}$ is in $k(\sqrt{r})$, for some $s \neq 0$ in k*
3. *Either $\sqrt{r(b^2r-a^2)}$ or $\sqrt{a^2-b^2r}$ is in k.*

Finally, they showed that, for denesting expressions containing only real square roots, only square roots or fourth roots play a role:

THEOREM 3 [3]. *Let k be a real extension of Q, and let a, b, r be in k with $\sqrt{r}$ not in k. Let $n_1, \ldots, n_l \geq 1$, and let $a_1, \ldots, a_l$ in k be positive. If $\sqrt{a+b\sqrt{r}}$ is in $k(\sqrt[n_1]{a_1}, \ldots, \sqrt[n_l]{a_l})$, then $\sqrt{a+b\sqrt{r}}$ is in $k(\sqrt[4]{r}, \sqrt{p_1}, \ldots, \sqrt{p_l})$ for some $p_1, \ldots, p_l$ in k.*

In combination, these three theorems provide the backbone for an algorithm for denesting real nested square roots. Let r_0 be in Q, and inductively let $r_i = a_i + b_i\sqrt{r_{i-1}}$, with a_i and b_i in $Q(r_{i-1})$ for $i = 1, \ldots, n$. Consider the following tower of fields:

$$k_0 = Q, \text{ and for all } i \geq 0, k_{i+1} = k_i(\sqrt{r_i}), \text{ where } r_i \in k_i.$$

There are two ways $\sqrt{r_n}$ can denest using only square roots. One is that $\sqrt{r_{n-1}}$ may denest. The other is that there is a field K satisfying the following (i) it contains a, b, r_{n-1}, (ii) it contains only elements of depth $n-1$, and (iii) $a^2 - b^2 r_{n-1}$ is a square.

To find the field K in which these conditions are satisfied, I attempt to denest $\sqrt{a^2 - b^2 r_{n-1}}$. If this is accomplished, I will have found a field $\hat{k}$ in which all elements have depth at most $n-2$. I will also have found an element s such that $\sqrt{a^2 - b^2 r_{n-1}}$ is in $\hat{k}(\sqrt{s}, \sqrt{r_{n-2}})$. Then $K = \hat{k}(\sqrt{s}, \sqrt{r_{n-2}})$. This idea leads to a recursive algorithm. If one is careful with the input and output, the running time of this algorithm is polynomial.

But this tells us only how to handle a nested square root which consists of roots recursively formed by $r_i = a_i + b_i\sqrt{r_{i-1}}$, with a_i and b_i in $Q(r_{i-1})$ for $i = 1, \ldots, n$, and r_0 in Q. If we want to denest all real nested square roots, we have to be able to handle linear combinations of nested square roots. I look first at the simpler case, in which none of the radicals is nested.

THEOREM 4 (Besicovitch [2]). *Let $\{e_i\}$ denote the set of n^l radicals,*

$$\sqrt[n]{p_1^{m_1} p_2^{m_2} \cdots p_l^{m_l}}, \quad 0 \leq m_i \leq n, \quad 1 \leq i \leq l,$$

where $p_1, \ldots, p_l$ are the first l primes. Then the set $\{e_i\}$ is linearly independent over Q.

To check if a linear combination of square roots is equal to zero, one needs only to check if it is trivially equal to zero. There is the obvious solution of factoring all the integers under the square root sign in order to check if pieces cancel, but this is far from optimal. A much faster way to do the problem is to compute gcd's of the integers under the square root signs. Then where appropriate, pull out squares from under the radical signs. Combine like terms. Repeat until no further simplification is possible. If there is anything left, then the combination of square roots is different from 0.

One can implement a more complex version of this idea to simplify linear combinations of nested square roots. Borodin, *et al.* [3] consider the case where k is a real extension of Q, and where l_i, a_i, b_i, and $\sqrt{r_i}$ are in k for $i = 1, \ldots, h$. Suppose $\sum_{i=1}^{h} l_i\sqrt{a_i + b_i\sqrt{r_i}}$ denests in k. They find the denesting as follows. First denest any single radical that can be denested using the criteria of Theorem 3. Next consider each pair of nested radicals, $\sqrt{a_i + b_i\sqrt{r_i}}, \sqrt{a_j + b_j\sqrt{r_j}}$, and see if the product denests. Suppose it is equal to m in k. Replace $l_i\sqrt{a_i + b_i\sqrt{r_i}} + l_j\sqrt{a_j + b_j\sqrt{r_j}}$ by $[l_i + ml_j/(a_i + b_i\sqrt{r_i})] \cdot \sqrt{a_i + b_i\sqrt{r_i}}$ in $k(\sqrt{a_i + b_i\sqrt{r_i}})$. Iterate the process of looking for a pair of radicals that denests. If at any point the product of any pair of radicals cannot be further denested, then the combination of nested radicals cannot be further denested.

Earlier Siegel had studied a more general situation than square roots. Assume that F is an arbitrary real field, and let $r_1, \ldots, r_k$ be natural numbers. Let $q_1, \ldots, q_k$ be elements of F, with $\sqrt[r_1]{q_1}, \ldots, \sqrt[r_k]{q_k}$ real. Siegel introduced the multiplicative groups generated by the nonzero elements of F and the first i radicals $\{\sqrt[r_1]{q_1}, \ldots, \sqrt[r_i]{q_i}\}$, which he denoted by $\Gamma^{(i)}$. Let r_i' be the group index of $\Gamma^{(i-1)}$ in $\Gamma^{(i)}$. Then Siegel showed:

THEOREM 5 [12]. *With notation as above, the degree of the field extension $F(\sqrt[r_1]{q_1}, \ldots, \sqrt[r_k]{q_k})$ equals $\prod_{i=1}^{k} r_i'$. Thus, the basis of the extension is given by*

$$\prod_{i=1}^{k} \sqrt[r_i]{\beta_i^{e_i}}, \quad 0 \leq e_i \leq r_i' - 1, \, i = 1, \ldots, k.$$

In seeking an algorithm for simplifying sums of radicals. Blömer recast this as

COROLLARY 6 [4]. *Let $S = \sum_{i=1}^{k} \gamma_i \sqrt[r_i]{\beta_i}$, where the r_i are natural numbers, and γ_i, β_i are in F with $\sqrt[r_i]{\beta_i}$ real. Assume that there is no pair of indices $(i, j), i \neq j$, with $\sqrt[r_i]{\beta_i}/\sqrt[r_j]{\beta_j} \in F$. Then S is in F if and only if there exists at most one $\gamma_i \neq 0$. In this case, $\sqrt[r_i]{\beta_i}$ is in F. That is, arbitrary real radicals $\sqrt[r_i]{\beta_i}$ are linearly dependent over F if and only if there are already two radicals $\sqrt[r_i]{\beta_i}, \sqrt[r_j]{\beta_j}$ that are linearly dependent over F.*

This gives an easy algorithm to check if a sum of radicals over a field $Q(\alpha)$ is in $Q(\alpha)$. Let $F = Q(\alpha)$ be an algebraic number field, let $S = \sum_{i=1}^{k} \gamma_i \sqrt[r_i]{\beta_i}$, and let $R = \{\sqrt[r_1]{\beta_1}, \ldots, \sqrt[r_k]{\beta_k}\}$, where the β_i are in F. To show that $\sqrt[r_1]{\beta_1}/\sqrt[r_2]{\beta_2}$ is in F, it is sufficient to show that

1. $\sqrt[r_1']{\beta_1} = \beta_1' \in F$,
2. $\sqrt[r_2']{\beta_2} = \beta_2' \in F$,
3. $\sqrt[r_1']{\beta_1'}/\sqrt[r_2']{\beta_2'} \in F$,

where $r = \gcd(r_1, r_2)$ and $r_1' = r_1/r$, $r_2' = r_2/r$. If $F = Q$ and $q = z_1/z_2$, with $\gcd(z_1, z_2) = 1$, clearly $\sqrt[r]{q}$ is in Q if and only if $\sqrt[r]{z_1}$ and $\sqrt[r]{z_2}$ are in Z. Blömer uses logarithms to compute an approximation to $\sqrt[r]{z}$, which is then checked to see if it is an rth root of z. If $F \neq Q$, Blömer [4] gives a Monte Carlo[2] algorithm with error probability 2^{-t} which, given α, β, and γ, decides in time polynomial in $\log r$ whether there is a γ in $Q(\alpha)$ that satisfies $\gamma^r = \beta$. He does this by checking whether γ is an rth root of β modulo T distinct primes which are randomly chosen

[2] Monte Carlo and Las Vegas methods are the most common types of probabilistic algorithms used in computer science. The running time is polynomial in Monte Carlo algorithms, with an exponentially small chance of failure. Las Vegas algorithms always produce the correct answer with a polynomial expected running time.

from the interval $[0, 2^T]$, where T is a polynomial in $\log r$, the input size of β, and t. Blömer's technique also explains the denestings of the radicals at the beginning of this article.

Call depth 2 nested real radicals "Ramanujan Radicals." Building on Theorem 5, Blömer proved

THEOREM 7 [5]. *Let F be a real field, and $F' = F(\sqrt[r_1]{q_1}, \ldots, \sqrt[r_k]{q_k})$ be a real radical extension of F of degree r. Let γ be in F', and let s be a natural number. Assume that $\sqrt[s]{\gamma}$ is real and denests over F. Then there is a nonzero η in F such that $\sqrt[sr]{\eta}\sqrt[s]{\gamma}$ is in F'.*

Consider the field embeddings σ_j of $F(\sqrt[r_1]{\beta_1}, \ldots, \sqrt[r_k]{\beta_k})$ in its splitting field over F. Let $\zeta_{r_i'}$ be a primitive r_i'th root of unity, where r_i' is the group index of $\Gamma^{(i-1)}$ in $\Gamma^{(i)}$. Let $0 \leq f_i \leq r_i'$. The field embeddings σ_j are given by

$$\sigma_j : F(\sqrt[r_1]{\beta_1}, \ldots, \sqrt[r_k]{\beta_k}) \to F(\zeta_{r_1'}^{f_1} \sqrt[r_1]{\beta_1}, \ldots, \zeta_{r_k'}^{f_k} \sqrt[r_k]{\beta_k}).$$

Then

THEOREM 8. *Assume $\sqrt[s]{\gamma}$ denests as above. Let $\{\beta_i\}$ be a basis of F' over F. Then for some basis element β_i, $\sum_{j=1}^{r} \sigma_j(\beta_i) \sqrt[s]{\sigma_j(\gamma)}$ is a nonzero element of F and $[\sum_{j=1}^{r} \sigma_j(\beta_i) \sqrt[s]{\sigma_j(\gamma)}]^{-rs}$ denests $\sqrt[s]{\gamma}$.*

If we let $\eta = [\sum_{j=1}^{r} \sigma_j(\beta_i)(\sqrt[s]{\sigma_j(\gamma)})]^{-rs}$, then η denests $\sqrt[s]{\gamma}$, that is, $\sqrt[rs]{\eta}\sqrt[s]{\gamma} = \gamma' \in F'$. The value of this theorem is that in bounding η, it tells us where to search for a denesting.

Let us consider the simplest possible case, a Ramanujan Radical. The base field $F = Q$, and $F' = Q(\sqrt[r_1]{q_1}, \ldots, \sqrt[r_k]{q_k})$ with $r_i \in N$ and $q_i \in Q$. Interpret $\sqrt[r_i]{q_i}$ to be the real r_ith root. I wish to denest $\sqrt[s]{\gamma}$, where $\gamma = \sum_{i=1}^{k} d_i\beta_i$, and the $\{\beta_i\}$ form the basis of F' over F, and the d_i are in Q. Without loss of generality, one can assume that the d_i and q_j are all integers, that is, that γ is an algebraic integer.

If γ is an algebraic integer, so is $\sigma_j(\gamma^r)$ for each j. Furthermore, $(\sigma_j(\gamma^r))^{-rs}$ is one of the rs roots of $\sigma_j(\gamma^r)$. Thus, $(\sigma_j(\gamma^r))^{-rs}$ is an algebraic integer, and so is $[\sum_{j=1}^{r} \sigma_j(\beta_i) \sqrt[s]{\sigma_j(\gamma)}]^{-rs}$, and β_i is an element of any basis for F' over F for which $\beta_i\gamma'$ has nonzero trace. As before, Blömer employs the idea of computing logs and lattice reduction to compute η.

If the sum denests, Theorem 8 tells us how. This denesting will have depth 1, that is, it will be a sum of radicals. We can easily check whether this sum is actually in Q or not. Thus, Theorem 8 leads to an algorithm for completely denesting Ramanujan Radicals. Note that Blömer's technique guarantees a minimal-depth solution only for the depth 2 case, Ramanujan Radicals. Blömer also has a nice solution to the question of sums of nested radicals:

THEOREM 9 [5]. *Suppose $S = \sum_{i=1}^{k} \beta_i \sqrt[s_i]{\gamma_i}$ is a sum of real nested radicals such that γ_i is in F_i, $\beta_i \in L_i$, $i = 1, \ldots, k$, where each F_i, L_i is a real radical extension of F, a subfield of the reals of degree r_i, t_i, respectively. Assume that no nested radical $\sqrt[r_i]{\gamma_i}$ denests using real radicals. Then*

1. *S can denest only to 0;*
2. *if no quotient $\sqrt[r_i]{\gamma_i} / \sqrt[r_j]{\gamma_j}$ denests, then $S = 0$ if and only if $\beta_i = 0$ for all i.*

This gives a polynomial-time algorithm to determine whether the sum of nested real radicals is equal to zero.

A Particularly Simple Form of Denesting

Zippel studied the equation

$$\sqrt[3]{\sqrt[3]{2} - 1} = \sqrt[3]{1/9} - \sqrt[3]{2/9} + \sqrt[3]{4/9}$$

and found the beginnings of a pattern. Whereas $\sqrt[3]{2} - 1$ is not a cube in $Q(\sqrt[3]{2})$, $9(\sqrt[3]{2} - 1)$ is. I can find a γ which satisfies $\gamma^3 = 9(\sqrt[3]{2} - 1)$ in $Q(\sqrt[3]{2})$ by factoring $x^3 - 9(\sqrt[3]{2} - 1)$ over $Q(\sqrt[3]{2})$. Namely,

$$\begin{aligned} &x^3 - 9(\sqrt[3]{2} - 1) \\ &\quad = (x - (1 - \sqrt[3]{2} + \sqrt[3]{4}))(x^2 + (1 - \sqrt[3]{2} + \sqrt[3]{4})x + 3\sqrt[3]{4} - 3). \end{aligned}$$

The first factor gives Ramanujan's denesting.

Zippel noticed that a similar situation arises with $5 + 2\sqrt{6}$. Again, this is not a square in $Q(\sqrt{6})$, but a multiple of it, $2(5 + 2\sqrt{6})$, is. Our task is to find γ, where $\gamma^2 = 2(5 + 2\sqrt{6})$ in $Q(\sqrt{6})$. We discover

$$x^2 - 2(5 + \sqrt{6}) = (x - (2 + \sqrt{6}))(x - (-2 - \sqrt{6})).$$

In both cases, we have found an element β in Q such that, although $\sqrt[d]{\alpha}$ is not an element of $Q(\theta)$, $\sqrt[d]{\alpha\beta}$ is. We have the following picture:

$$\begin{tikzcd}
 & L = K(\sqrt[d]{\alpha}) = KF \arrow[dl, dash] \arrow[dr, dash] & \\
K \arrow[dr, dash] & & F = k(\sqrt[d]{\alpha}) \arrow[dl, dash] \\
 & k = K \cap F &
\end{tikzcd}$$

In each case, expressions of nesting depth n in the field L have been dropped to expressions of nesting depth $n - 1$ in the subfield L. This idea generalizes to a theorem:

THEOREM 10 (Zippel [13][3]). *Assume K is an extension of k, a field containing a primitive dth root of unity. Let $L = K(\sqrt[d]{\alpha})$ be an extension of degree d, where α is in K. If there is a field F which is a Galois extension of $k = K \cap F$, and $L = KF$, then there is a β in k such that $\alpha\beta$ is a dth power of an element of k. Furthermore, $F = k(\sqrt[d]{\beta})$.*

Zippel exploited some lucky guesses. If I want an algorithm, I will need something somewhat more deterministic than that. I will need an algorithm to determine whether such a β exists, and if so, how to find it. More precisely, given $\sqrt[d]{\alpha}$ in L, when is there a solution to $\alpha\beta = \gamma^d$,

[3] Zippel's original statement omitted, but implicitly assumed, the hypothesis that F is a Galois extension of k. A corrected version appears in [9].

with β in k, a proper subfield of $k(\sqrt[d]{\alpha})$, and γ in K, a proper subfield of L? It is not hard to show that the following converse of Zippel's theorem holds:

THEOREM 11 (Landau [9]). *Let α be an element of a field K. Suppose that $\sqrt[d]{\alpha}$ is of degree d over K, and that $\sqrt[d]{\alpha} = \lambda / \sqrt[d]{\beta}$, with λ in K and β in $k \subset K$. Assume that the* dth *roots of unity lie in k. Then the field $F = k(\sqrt[d]{\beta})$ satisfies:* (i) *F over k is Galois and the Galois group of F over $F \cap K$ is isomorphic to the group of FK over K,* (ii) *$FK = K(\sqrt[d]{\alpha})$, and* (iii) *$k = F \cap K$.*

Thus, if I seek a "Zippel denesting," I am asking to find subfields of $K(\sqrt[d]{\alpha})$ satisfying Theorem 11. In [10], there is a polynomial-time algorithm to find maximal subfields of a field. One can use this for an algorithm for determining whether a Zippel denesting exists. The first step is to find all maximal subfields of L.

It is a simple matter to check if a field extension is Galois. One finds a primitive element for the larger field, possibly by resorting to the well-known construction that $k(\gamma, \rho) = k(\gamma + c\rho)$ for some $c \leq (\deg_k(\gamma)\deg_k(\rho))^2$. With a primitive element, say α, which has minimal polynomial $p(x)$, one can compute the action of the Galois group by observing that the factorization

$$p(x) = (x - p_1(\alpha))(x - p_2(\alpha)) \cdots (x - p_m(\alpha))$$

gives the group table, since $\sigma_i(\alpha_j) = (p_i(p_j(\alpha))$ (mod $p(x)$).

Computing whether a candidate subfield satisfies parts (ii) and (iii) is even easier. It is just a matter of checking whether two fields are equal. This can be done by seeing whether the basis of one is contained in the other, and vice versa. Iterating the procedure in [10] will give an algorithm to find all subfields. Thus, there is an algorithm to determine if a "Zippel denesting" exists.

It is not fast. There may be 2^d fields between k and L, and in a worst case I would have to check each one of them. But this exponential-time algorithm can still be quite reasonable for small (< 10) values of d.

Zippel used his theorem to shed some light on the calculations and theorems of Borodin, *et al.* [3]. Let a, b, and q be elements of a field k, and suppose I am hoping to denest $\sqrt{a + b\sqrt{q}}$. Assume there is a β in k such that

$$\beta(a + b\sqrt{q}) = (a_0 + \sqrt{q})^2.$$

Now $a^2 - qb^2$, the norm of $a + b\sqrt{q}$, is a square, d^2. This leads to

$$\beta = \frac{2}{b^2}(a \pm d), \qquad a_0 = \frac{1}{b}(a \pm d).$$

Choosing the positive sign,

$$\sqrt{a + b\sqrt{d}} = \sqrt{\frac{a+d}{2}} \pm \sqrt{\frac{a-d}{2}},$$

where the sign depends on the sign of b.

If $a^2 - qb^2$ is not a square, then that means that $\sqrt{a + b\sqrt{q}}$ does not denest in a quadratic extension $K = k(\sqrt{q})$, and I must look for a quartic extension in which it denests. I try looking for a β such that $\beta(a\sqrt{q} + bq)$ is a perfect square. That computation eventually leads to

$$\sqrt{a + b\sqrt{q}} = \frac{1}{2(bq + d)}\left(\sqrt[4]{4q(bq + d)^2} + \sqrt[4]{(4q(bq + d)^2)^3}\right).$$

These two denesting formulas are the two shown in [3] to be the only ways in which expressions involving square roots can be denested.

What About the General Case?

Zippel's criteria are simple and elegant, but the conditions in Theorems 10 and 11 are sufficiently restrictive that they will not handle all cases. Trying to understand where the 9 came from in

$$\sqrt[3]{\sqrt[3]{2} - 1} = \sqrt[3]{1/9} - \sqrt[3]{2/9} + \sqrt[3]{4/9},$$

I discovered the following subfields[4] of $Q(\sqrt[3]{\sqrt[3]{2} - 1})$:

$$\begin{tikzcd}
 & Q(\sqrt[3]{\sqrt[3]{2} - 1}) = Q(\sqrt[3]{2}, \sqrt[3]{9}) \arrow[dl, dash] \arrow[dr, dash] & \\
Q(\sqrt[3]{2}) \arrow[dr, dash] & & Q(\sqrt[3]{9}) \arrow[dl, dash] \\
 & Q &
\end{tikzcd}$$

The denesting

$$\sqrt[6]{7\sqrt[3]{20} - 19} = \sqrt[3]{5/3} - \sqrt[3]{2/3}$$

led to a similar tower of fields:

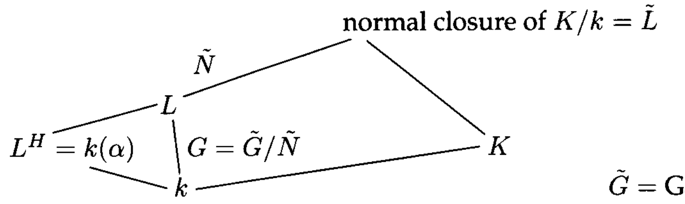
$$\begin{tikzcd}
 & Q\left(\sqrt[6]{7\sqrt[3]{20} - 19}\right) = Q(\sqrt[3]{20}, \sqrt[3]{18}) \arrow[dl, dash] \arrow[d, dash] \arrow[dr, dash] & \\
Q(\sqrt[3]{20}) \arrow[dr, dash] & Q(\sqrt[3]{18}) \arrow[d, dash] & Q(\sqrt[3]{45}) \arrow[dl, dash] \\
 & Q &
\end{tikzcd}$$

In fact, such pictures arose for *all* the simplifications I had occasion to try. This was too beautiful to happen by accident. A natural place to search for a denesting is the smallest closed field in which the minimal polynomial of α factors completely — the splitting field. The answer almost turned out to be that surprisingly simple and elegant. A minimal-depth expression for a nested radical can always be found in the splitting field, *provided all roots of unity lie in the base field.* More precisely:

THEOREM 12 (Landau [8]). *Suppose α is a nested radical over k, where k is a field of characteristic 0 containing all roots of unity. Then there is a minimal-depth nesting of α with each of its terms lying in the splitting field of the minimal polynomial of α over k.*

[4] This diagram gives a partial explanation of where the 9 comes from. A more complete answer is that 9 divides the discriminant of $x^9 + x^6 + x^3 - 1$, the minimal polynomial of $\sqrt[3]{\sqrt[3]{2} - 1}$ over Q.

Proof. Consider the following diagram:


normal closure of $K/k = \tilde{L}$
$\tilde{N}$
L
$L^H = k(\alpha)$
$G = \tilde{G}/\tilde{N}$
K
k

$\tilde{G}$ = Galois group of $\tilde{L}/k$

Let α be a nested radical over k, and suppose that L is the splitting field of $k(\alpha)$ over k. Let $\tilde{L}$ be the normal closure of K over k, where K is a field which contains a minimal-depth nested expression for α. If I let $\tilde{G}$ be the Galois group of $\tilde{L}$ over k, and G be the group of L over k, then I note that $G = \tilde{G}/\tilde{N}$, where $\tilde{N}$ is the group of $\tilde{L}$ over L. (Note that $\tilde{N}$ is normal because $\tilde{L}$ is normal over L.) As α can be denested over L, there is a sequence of subgroups $\tilde{H}_1, \ldots, \tilde{H}_l$ of $\tilde{G}$ with $\tilde{G} = \tilde{H}_0 \triangleright \tilde{H}_1 \triangleright \cdots \triangleright \tilde{H}_l$, with $\tilde{H}_i/\tilde{H}_{i+1}$ abelian for $i = 0, \ldots, l-1$, and $\tilde{H} \supset \tilde{H}_l$. This sequence can be pulled down to a sequence in G. If all roots of unity are in k, the tower defined by the groups can be made into a tower of radical extensions, thus showing there is a minimal-depth denesting in L. ∎

Of course, this does not solve the original problem, which was to denest over an arbitrary field. Can one *a priori* add certain roots of unity to the base field so that a minimal-depth nesting can be achieved? The answer is yes, so long as one is careful in handling roots of unity.

All roots of unity can be expressed in terms of radicals. The problem is that the depth of nesting of a root of unity can be very deep indeed. In general, a pth root of unity has nesting depth one more than the maximum of the nesting depths of the prime factors of $p-1$. Thus, if there are arbitrarily long sequences of primes $p, 2p+1, 2(2p+1)+1, \ldots$ — a plausible, but unproved conjecture in number theory — then an nth root of unity can have nesting depth $\log n$.

For me the motivation for studying the denesting of radicals was to develop an algorithm for radical simplifications. In many applications, writing a root of unity as ζ_n instead of the nested radical is a perfectly reasonable solution. This was the route taken here. But adding roots of unity to k does change the field in unexpected ways.

By the Kronecker–Weber Theorem, every abelian extension over Q can be embedded in a cyclotomic extension. When I attempt to write $\sqrt[n]{\alpha}$ in $Q(\zeta_l)$ I may find that $\sqrt[n]{\alpha}$ is an irrational number which is already in $Q(\zeta_l)$. Such is the case for $\sqrt{5}$ in the field $Q(\zeta_5)$. Thus, $\sqrt{5}$ will be represented as a polynomial in ζ_5, rather than the more usual expression $\sqrt{5}$. This type of simplification may drop us a single level of nesting. A more serious problem is that writing a root of unity as ζ_l in some sense masks it. There are subtle ways in which I pay for that. For example, $\sqrt{\sqrt{5}-5/2} = \zeta_5 - 1/\zeta_5$. Which symbol is easier to understand: $\sqrt{\sqrt{5}-5/2}$ or $\zeta_5 - 1/\zeta_5$? That depends on the application — or the mathematician.

Taking these concerns into consideration, I find:

THEOREM 13 (Landau [8]). *Suppose α is a nested radical over k, where k is a field of characteristic 0. Let L be the splitting field of $k(\alpha)$ over k, with Galois group G. Let l be the lcm of the exponents of the derived series of G, and let us write a primitive lth root of unity as ζ_l, and not simply as a nested radical. If there is a denesting of α such that each of the terms has depth no more than t, then there is a denesting of α over $k(\zeta_l)$ with each of the terms having depth no more than $t+1$ and lying in $L(\zeta_l)$.*

I also have an alternative version of this result in which I achieve minimal depth at the expense of adjoining a primitive rth root of unity, where r is dependent on the presentation of the input.

COROLLARY 14 [8]. *Let k, α, L, G, l, and t be as in Theorem 13. Let m be the lcm of the (m_{ij}), where m_{ij} runs over all the roots appearing in the given nested expression for α. Let r be the lcm of (m, l). Then there is a minimal-depth nesting of α over $k(\zeta_r)$ with each of its terms lying in $L(\zeta_r)$.*

These theorems tell us that the splitting field is the right place to look. They also lead naturally to an algorithm. If I wish to denest the nested radical α, I begin by computing the minimal polynomial of α over k. From that I construct the splitting field L of the minimal polynomial of α over k. Next I compute $G = \mathrm{Gal}(L/k)$. We have already seen how to do these computations. What is the shortest sequence of nested radicals that will give α? It will come from the shortest sequence of groups in the Galois group, the series of commutator subgroups $D^iG, i = 1, \ldots, s$, where $D^sG = \{e\}$. Good algorithms for group computations have existed for over 15 years. Having a group table, or an equivalent, for G, it is not hard to compute the commutator series of the group.

Next I also compute l, the lcm of the exponents of the derived series of G. For each $i, i = 1, \ldots, s$, I compute $D^{i-1}G/D^iG = J_{i1} \times \cdots \times J_{it_i}$ as a direct product of cyclic groups. Let $\tilde{J}_{ij} = \{e\} \times \cdots \times \{e\} \times J_{ij} \times \{e\} \times \cdots \times \{e\}$, and let $L_i = L^{D^iG}$. Thus, for each $i, L_i = L_i^{\tilde{J}_{i1}} \cdots L_i^{\tilde{J}_{it_i}}$ is a composite of cyclic extensions of L_{i-1}. For each i and j, I compute $\tilde{\beta}_{ij}$ such that $L_i^{\tilde{J}_{ij}} = L_{i-1}(\tilde{\beta}_{ij})$. Thus, $L_i = L_{i-1}(\tilde{\beta}_{i1}, \ldots, \tilde{\beta}_{it_i})$.

I write $K_0 = k(\zeta_l)$, where ζ_l is a primitive lth root of unity. Then $K_{ij} = K_{i-1}(\tilde{\beta}_{ij})$ can be written as a radical extension of K_{i-1}, and each $K_i = K_{i1} \cdots K_{it_i}$ is a composite of radical extensions of K_{i-1}. The crux of the matter is how to write these extensions as radical ones, that is, $K_i = K_{i-1}(\sqrt[d_i]{\alpha_{i-1}})$. This is achieved as follows.

Following Artin, I construct a polynomial $s_{ij}(x)$ whose roots $\theta_{ij1}, \ldots, \theta_{ijr_{ij}}$ form a "normal" basis for K_{ij} over K_{i-1}. The degree of $s_{ij}(x)$ is $r_{ij} = [K_{ij} : K_{i-1}]$, and its roots are linearly independent over K_{i-1}. Then I use Lagrange resolvents to find a β_{ij} in K_{ij} such that $K_{ij} = K_{i-1}(\beta_{ij})$, where β_{ij} satisfies an irreducible polynomial of the form $x^{n_{ij}} - b_{ij}$ over K_{i-1}. That each of the extensions can be obtained as radical extensions stems from the fact that the appropriate roots of unity lie in the base field.

This is just a brief sketch of the algorithm, details of which can be found in [8]. But the point should be clear: There is an algorithm for simplifying nested radicals, assuming one allows roots of unity to creep into the expression.

How long does this take? Too long! If α is of degree n over k, its Galois group may be of size $n!$ Even groups which are exponentially large (S_n, A_n, etc.) have a small set of generators, but from a computational standpoint, that does not seem to help. No one knows how to determine the generators of a Galois group of a general polynomial without first determining its splitting field. In general, computing the splitting field (that is, finding a minimal polynomial for a generator over the base field) is an exponential-time computation. Allowing roots of unity written in shorthand, Theorem 12 limits where I have to search if I am seeking a denesting. Nonetheless, except for radicals with small degrees, the computation is presently infeasible. Until there are improvements in algorithms for splitting field and Galois group computations, the algorithms based on Theorems 12, 13, and Corollary 14 are useful only for nested radicals of small degree.

Problem 1: Find a polynomial-time algorithm to compute the Galois group of an irreducible polynomial over Q.

There is an improvement one can make to Corollary 14. Horng and Huang [7] have shown:

THEOREM 15 [7]. *Let k, α, L, G, l, and t be as in Theorem 13. Let n be a natural number which is divisible by $[L : k]$ and the discriminant of L over Q. Then there is a minimal-depth nesting of α over $k(\zeta_n)$ with each of its terms lying in $L(\zeta_n)$.*

In finding a root of unity to achieve minimal-depth nesting for α, they eliminate the need for including anything that relies on the presentation of α, as in Corollary 14. However, they do so at the expense of introducing a dth root of unity, where the discriminant of L over Q divides d. This discriminant is of exponential size in α. Their algorithm for denesting follows the algorithm described earlier.

The problem of simplification of nested radicals is far from completely solved. Because of Blömer we now can efficiently denest depth-2 real radicals. We do not have efficient algorithms for depth 3 or greater real nested radicals which achieve minimal-depth nestings. Thus, the following questions remain:

Problem 2: Without a special encoding for roots of unity, given a nested radical, determine whether there is another nested radical of the same value, with lower nesting depth.

Problem 3: Find such a lower nesting-depth radical.

Problem 4: Given a real nested radical, determine whether there is another nested radical of the same value, of lower nesting depth.

Acknowledgments

I would like to thank Neil Immerman, Hendrik Lenstra, Carl Pomerance, Martin Tompa, and Ann Trenk for their help.

References

1. E. Artin, *Galois Theory*, University of Notre Dame Press, Notre Dame, IN, 1942.
2. A. Besicovitch, On the linear independence of fractional powers of integers, *J. London Math. Soc.* 15 (1940), 3–6.
3. A. Borodin, R. Fagin, J. Hopcroft, and M. Tompa, Decreasing the nesting depth of expressions involving square roots, *J. Symb. Comput.* 1 (1985), 169–188.
4. J. Blömer, Computing sums of radicals in polynomial time, *Proceedings of the 32nd Annual IEEE Symposium on Foundations of Computer Science*, 1991, pp. 670–677.
5. J. Blömer, How to denest Ramanujan's nested radicals, *Proceedings of the 33rd Annual IEEE Symposium on Foundations of Computer Science*, 1992, pp. 447–456.
6. B. Caviness and R. Fateman, Simplification of radical expressions, *Proc. SYMSAC 77*, pp. 329–338.
7. G. Horng and M. Huang, On simplifying nested radicals and solving polynomials by pure nested radicals of minimum depth, *Proceedings of the 31st Annual IEEE Symposium on Foundations of Computer Science*, 1990, pp. 847–854.
8. S. Landau, Simplification of nested radicals, *SIAM J. Comput.* 21 (1992), 85–109.
9. S. Landau, A Note on Zippel denesting, *J. Symb. Comput.* 13 (1992), 41–45.
10. S. Landau and G. Miller, Solvability by radicals is in polynomial time, *J. Comput. Syst. Sci.* 30 (1985), 179–208.
11. S. Ramanujan, *Problems and Solutions, Collected Works of S. Ramanujan*, Cambridge University Press, Cambridge, 1927.
12. C. Siegel, Algebraische Abhängigkeit von' Wurzeln, *Acta Arithmetica* 21 (1971), 59–64.
13. R. Zippel, Simplification of expressions involving radicals, *J. Symb. Comput.* 1 (1985), 189–210.

Computer Science Department
University of Massachusetts
Amherst, MA 01003 USA
e-mail:landau@cs.umass.edu