Substitutions in Integrals; Some Pitfalls

Louis A. Talman
Department of Mathematical & Computer Sciences
Metropolitan State College of Denver

March 8, 2006

We commonly make substitutions involving common elementary functions that are not one-to-one, and this common practice leads us into making our substitutions in ways that violate a number of principles. This phenomenon typically evidences itself (or *should* evidence itself) as an unresolved ambiguity of sign arising from a trigonometric substitution.

Consider, for example, the integral $\int \frac{dx}{x^3\sqrt{x^2-1}}$. In common practice we deal with this indefinite integral in this way:

Make the substitution $x = \sec\theta$. Then dx becomes $\sec\theta\tan\theta\,d\theta$. Thus

$$\begin{aligned}
\int \frac{dx}{x^3\sqrt{x^2-1}} &= \int \frac{\sec\theta\tan\theta\,d\theta}{\sec^3\theta\sqrt{\sec^2\theta-1}} && (1)\\
&= \int \frac{\sec\theta\tan\theta\,d\theta}{\sec^3\theta\sqrt{\tan^2\theta}} && (2)\\
&= \int \frac{\sec\theta\tan\theta\,d\theta}{\sec^3\theta\tan\theta} && (3)\\
&= \int \cos^2\theta\,d\theta && (4)\\
&= \frac{1}{2}\int (1+\cos 2\theta)\,d\theta && (5)\\
&= \frac{1}{2}\left(\theta + \frac{1}{2}\sin 2\theta\right) + c && (6)\\
&= \frac{1}{2}(\theta + \sin\theta\cos\theta) + c. && (7)
\end{aligned}$$

But $\sec\theta = x$, and so $\theta = \operatorname{arcsec} x$ and $\cos\theta = \frac{1}{x}$, while

$$\sin\theta = \sqrt{1-\cos^2\theta} \qquad (8)$$

$$= \sqrt{1 - \frac{1}{x^2}} \tag{9}$$

$$= \sqrt{\frac{x^2 - 1}{x^2}} \tag{10}$$

$$= \frac{\sqrt{x^2 - 1}}{x}. \tag{11}$$

Consequently,

$$\int \frac{dx}{x^3\sqrt{x^2 - 1}} = \frac{1}{2}\left(\operatorname{arcsec} x + \frac{\sqrt{x^2 - 1}}{x^2}\right) + c. \tag{12}$$

Let us analyze this calculation carefully. The domain of the integrand is the set $\{\, x \in \mathbb{R} : |x| > 1 \,\}$, and the substitution $x = \sec\theta$ certainly appears to honor this domain, as $|\sec\theta| \geq 1$ for all θ. However, in passing from (2) to (3), we implicitly assume that $\tan\theta \geq 0$; that is, that θ lies in either the first quadrant or the third quadrant. This seems to cause no trouble, because $\sec\theta \geq 1$ when θ lies in the first quadrant while $\sec\theta \leq -1$ when θ lies in the third quadrant. Of course, the fact that an implicit assumption causes no trouble does not excuse us from the error of making that assumption *implicitly*; assumptions ought always to be *explicit*. And we ought to take warning from the fact that integrand on the right side of equation (4) is unfailingly non-negative—which is not the case for the original integrand.

Notice that there are two similar errors in the calculation involving equations (8) through (11). In writing (8), we have implicitly assumed that $\sin\theta \geq 0$. This places θ in the first or the second quadrant, and we are now in trouble. When we combine this assumption with the assumption we made earlier, we find that we have ruled out the possibility that $x < -1$—thus invalidating our calculation on a part of the domain of the integrand. We might take hope, nevertheless, in the fact that in passing from (10) to (11) we have used the implicit assumption that $x > 0$ in a way that cancels out the error of sign that crept in at equation (8): Equation (11) is correct whatever the sign of x! Of course, in taking such hope, we must abandon the principle that a calculation with several errors in it doesn't justify the conclusion it reaches.

There is still an insuperable difficulty, though—one which we can't ignore on the ground that we got the right answer. The calculation quite ignores possible differences in the definition of the inverse secant function. We can see what trouble this causes by imagining for the moment that we are among those who select the range of the inverse secant function to be $\left[0, \frac{\pi}{2}\right) \cup \left(\frac{\pi}{2}, \pi\right]$. For us, the ambiguities of sign that we have discussed above would then come together in altogether the wrong fashion, and we would have

$$\begin{aligned}
&\frac{d}{dx}\left[\frac{1}{2}\left(\operatorname{arcsec} x + \frac{\sqrt{x^2-1}}{x^2}\right) + c\right] \\
&\quad = \frac{1}{2}\left(\frac{1}{|x|\sqrt{x^2-1}} + \frac{2-x^2}{x^3\sqrt{x^2-1}}\right) && (13) \\
&\quad = \begin{cases} \dfrac{1}{x^3\sqrt{x^2-1}}, & \text{when } x > 1; \\[2ex] \dfrac{1-x^2}{x^3\sqrt{x^2-1}}, & \text{when } x < -1. \end{cases} && (14)
\end{aligned}$$

Thus, under these conditions, the substitution utterly fails to give the correct antiderivative on that half of the domain of the integrand for which $x < -1$.

Consider, on the other hand, the *definite* integral $\displaystyle\int_{-2}^{-\sqrt{2}} \frac{dx}{x^3\sqrt{x^2-1}}$. Correct application of the substitution theorem for definite integrals to the evaluation of this integral forces us to incorporate correct choices of ambiguous signs at every step. Thus, those who take the range of the inverse secant function to be $\left[0, \frac{\pi}{2}\right) \cup \left(\frac{\pi}{2}, \pi\right]$ as above must evaluate $\displaystyle\int_{-2}^{-\sqrt{2}} \frac{dx}{x^3\sqrt{x^2-1}}$ in the following way:

Put $f(x) = \dfrac{1}{x^3\sqrt{x^2-1}}$, and put $\varphi(\theta) = \sec\theta$. Then $-2 = \varphi\left(\dfrac{2\pi}{3}\right)$, while $-\sqrt{2} = \varphi\left(\dfrac{3\pi}{4}\right)$. By the substitution theorem,

$$\begin{aligned}
\int_{-2}^{-\sqrt{2}} \frac{dx}{x^3\sqrt{x^2-1}} &= \int_{\varphi(2\pi/3)}^{\varphi(3\pi/4)} f(x)\,dx && (15) \\
&= \int_{2\pi/3}^{3\pi/4} f[\varphi(\theta)]\,\varphi'(\theta)\,d\theta && (16) \\
&= \int_{2\pi/3}^{3\pi/4} \frac{\sec\theta\tan\theta\,d\theta}{\sec^3\theta\sqrt{\sec^2\theta-1}} && (17) \\
&= \int_{2\pi/3}^{3\pi/4} \frac{\sec\theta\tan\theta\,d\theta}{\sec^3\theta\sqrt{\tan^2\theta}}. && (18)
\end{aligned}$$

But $\dfrac{2\pi}{3} \le \theta \le \dfrac{3\pi}{4}$ in this integral, and we have $\sqrt{\tan^2\theta} = -\tan\theta$ for such θ.

Consequently

$$\int_{-2}^{-\sqrt{2}} \frac{dx}{x^3\sqrt{x^2-1}}$$

$$= -\int_{2\pi/3}^{3\pi/4} \cos^2\theta\, d\theta \tag{19}$$

$$= -\frac{1}{2}\int_{2\pi/3}^{3\pi/4} (1+\cos 2\theta)\, d\theta \tag{20}$$

$$= -\frac{1}{2}\left(\theta + \frac{1}{2}\sin 2\theta\right)\Big|_{2\pi/3}^{3\pi/4} \tag{21}$$

$$= -\frac{1}{2}\left(\frac{3\pi}{4} + \frac{1}{2}\sin\frac{3\pi}{2}\right) + \frac{1}{2}\left(\frac{2\pi}{3} + \frac{1}{2}\sin\frac{4\pi}{3}\right) \tag{22}$$

$$= \frac{2-\sqrt{3}}{8} - \frac{\pi}{24}. \tag{23}$$

On the other hand, one who chooses the range of the inverse secant function to be $\left[-\pi, -\frac{\pi}{2}\right) \cup \left[0, \frac{\pi}{2}\right)$ must do the calculation a different way:

Put $f(x) = \dfrac{1}{x^3\sqrt{x^2-1}}$, and put $\varphi(\theta) = \sec\theta$. Then $-2 = \varphi\left(\dfrac{-2\pi}{3}\right)$, while $-\sqrt{2} = \varphi\left(-\dfrac{3\pi}{4}\right)$. By the substitution theorem,

$$\int_{-2}^{-\sqrt{2}} \frac{dx}{x^3\sqrt{x^2-1}} = \int_{\varphi(-2\pi/3)}^{\varphi(-3\pi/4)} f(x)\, dx \tag{24}$$

$$= \int_{-2\pi/3}^{-3\pi/4} f[\varphi(\theta)]\,\varphi'(\theta)\, d\theta \tag{25}$$

$$= \int_{-2\pi/3}^{-3\pi/4} \frac{\sec\theta\tan\theta\, d\theta}{\sec^3\theta\sqrt{\sec^2\theta - 1}} \tag{26}$$

$$= \int_{-2\pi/3}^{-3\pi/4} \frac{\sec\theta\tan\theta\, d\theta}{\sec^3\theta\sqrt{\tan^2\theta}}. \tag{27}$$

But $\dfrac{-3\pi}{4} \le \theta \le \dfrac{-2\pi}{3}$ in this integral, and we have $\sqrt{\tan^2\theta} = \tan\theta$ for such θ. Consequently

$$\int_{-2}^{-\sqrt{2}} \frac{dx}{x^3\sqrt{x^2-1}}$$

$$= \int_{-2\pi/3}^{-3\pi/4} \cos^2 \theta \, d\theta \tag{28}$$

$$= \frac{1}{2} \int_{-2\pi/3}^{-3\pi/4} (1 + \cos 2\theta) \, d\theta \tag{29}$$

$$= \frac{1}{2} \left(\theta + \frac{1}{2} \sin 2\theta \right) \Bigg|_{-2\pi/3}^{-3\pi/4} \tag{30}$$

$$= \frac{1}{2} \left[-\frac{3\pi}{4} + \frac{1}{2} \sin \left(-\frac{3\pi}{2} \right) \right] - \frac{1}{2} \left[-\frac{2\pi}{3} + \frac{1}{2} \sin \left(-\frac{4\pi}{3} \right) \right] \tag{31}$$

$$= \frac{2 - \sqrt{3}}{8} - \frac{\pi}{24}. \tag{32}$$