Russian America: Canvassing the State Ticket
Political Cartoon, Unknown (circa 1867)
Granger Collection (granger.com)
1  Scott Stevens ’16
Prophets and Pilgrimages: A Comparison in Haruki Murakami’s Works

15  Caroline Wheeler ’14
Conceptions of Conception: Definitions of the Beginning of Life and their Effect on Abortion Regulation

43  Katherine Keller ’15
A Faulty Cause and Effect: Exploring the Fallacy of the Connection Between Dred Scott and the 14th Amendment

59  Joseph Klein ’15
Where Practicality Meets Idealism: Reasons for the Criticism of the Alaska Purchase

77  Emma LaPorte ’14 and Darren Mei ’14
Project Mind Control

105  Sam Redmond ’14
Singing Tesla Coil: Building a Musically Modulated Lighting Machine
Prophets and Pilgrimages: A Comparison in Haruki Murakami’s Works

Scott Stevens

Land of ground thundering, land of Pacific flames, Japan always seems to be forging unity and bolstering its national identity through disasters. Examples range from the ancient, like when the kamikaze, or “divine wind” of a typhoon sank a 140,000-man Mongol fleet in 1281, to the recent devastating 6.8-magnitude Kobe Earthquake in 1995. Japan has been devastated by nuclear reactor meltdowns, as with the 2011 Fukushima Daiichi disaster, and by religious cults’ terrorism operations when the cult Aum Shinrikyou injured 6,252 people in their 1995 Tokyo subway sarin gas attack. Enter Haruki Murakami, a popular contemporary Japanese author, who often dabbles in the surreal and magical realist techniques. Although the story content and levels of magical realism vary, Haruki Murakami’s short story collection after the quake and newest novel Colorless Tsukuru Tazaki and His Years of Pilgrimage, written after the Fukushima disaster, propose the same reflection for Japan and the rest of a troubled world. Murakami convinces us that disasters, albeit senseless and random, are rejuvenating. They shake up or distort mundane thinking and sharpen our minds, which strengthens us to chase after the realities we truly want to live in. Specifically in these two works, his brand of strange magical realism perfectly matches the absurdity in disaster, and presents personal and environmental vagaries as opportunities for “seer” archetypes to swoop in and offer the main characters a “pilgrimage” to a more fulfilling life.

In the afterword to the short story collection after the quake, Murakami talks about his inspiration from the Kobe Earthquake in January 1995 and the subway gas attack that March. He believes that, although nature caused one disaster, and humans caused the other, they both

This paper was originally written in Japanese for an Independent Study in Japanese language with Kumiko Sakamoto in the spring of 2014.
ultimately capsized Japan. “We, on the whole, believe that the ground we step on is unshakeable. We may even each take it as a ‘self-evident truth.’ However, that can suddenly liquefy under our feet... those invisible lethal weapons do not distinguish commuters in their attacks” (263). This short story collection was also his first major work featuring third-person narration. The change from first person emphasizes the main characters’ choices not only as part of a personal journey, but as part of external forces at work. All of the short stories have a main character whom the Kobe Earthquake affects and a “seer” character who offers him control over his fate through a journey of sorts, one that reimagines the character’s pain as a strength and allows him to direct his life towards fulfillment. We’ll start with the short story that gives the collection its Japanese name, “All of God’s Children Dance.”

A reveling youth named Yoshiya is our protagonist who wakes up to a hangover one morning, a mere side effect of the profound transformation he’s just undergone. Flashback to Yoshiya’s conception: Yoshiya’s mother, after “lovelessly fornicating with many men”(94), in high school, dated a doctor with a missing earlobe, and, despite careful contraceptives, bore his child. The doctor abandoned her, and she found faith in a religious sect whose members saw Yoshiya’s as a virgin birth, which meant that he was the Son of God. Growing up, Yoshiya struggled to improve at his passion, baseball. Yoshiya couldn’t believe that his father was God if his father wouldn’t answer his prayers and help him become a better baseball player, and so he gave up his religion. After the Kobe earthquake strikes, his mother treks to Western Japan to work as an aid volunteer with her fellow church members. Yoshiya’s response to these events is to ditch work and follow a man in a train with a missing earlobe, believing him to be his estranged biological father. He follows him through an industrial area with glaring mercury streetlights, with which Murakami’s magical realist language works perfectly: “an imaginary scene with no sign of human life around him, looking all together like it was thrown together in a dream” (101). Yoshiya tracks him down near an enormous, “unfriendly” concrete fence with barbed wire “thickly surrounding the top of it as if daring the world” containing a scrapyard with heaps of cars. The man goes into a dark al-
leyway, and “although Yoshiya was a little hesitant, he still followed the man and set his foot into the twilight”(102). This screams of a hero’s journey. When Yoshiya wriggles through a hole at the end of the alley and finds himself in a baseball field, the man nowhere to be found, we realize that this typically magical realist “unexplainable” disappearance reveals the search for his “father” as a pretense for Yoshiya to arrive at the destined field. Alone there, Yoshiya remembers his relationships so far in his life, all of which seem to have been essentially sensual and centered around dancing. He recalls what a good dancer he has become through dancing with these partners as he dances by himself in the field. Feeling connected with the world, he becomes aware of his potential as a force for good, a realization parallel to his mother’s experience of sexuality and its deeper significance. For his mother, it was giving birth; for Yoshiya, it is club dancing. Both of these lead them to find God. His mother found religion, and at the end, Yoshiya whispers, “God.” The hope and possible salvation in Yoshiya wouldn’t have occurred had the Kobe earthquake not happened, shaking his connections to his mother, his father, and to his sense of purpose.

In another short story, “Superfrog Saves Tokyo,” a human-sized frog enlists everyday loan collector, Katagiri, to stop a massive subterranean worm from destroying Tokyo as it did Kobe. What makes the magical realism so potent is the polite and non-assuming way the frog, Kaeru, talks to Katagiri: “I ought to have made an appointment before coming...”(153). Strange, but it charms the reader as it charms Katagiri. Katagiri, like most of Murakami’s characters, is a loner who does his best to deal with his humdrum days of collecting loans in the midst of Chinese mafiosos and the Yakuza. His life reflects post-bubble Japan’s economic and spiritual depression. However, Kaeru comes as the “seer” character to lead Katagiri to new meaning by battling the worm. On his way to their meeting place, Katagiri gets shot unconscious, only waking up in a hospital to find Kaeru congratulating him on their battle, which happened in a dream world underground. Kaeru says that the worm wounded him and that he feels an “anti-self” (165), the entropy in his heart, overtaking him, so he rests but is then consumed by a swarm of bugs that envelop the whole hospital room. Katagiri
wakes up and is distraught by Kaeru’s death, but takes strength from it and realizes he can battle his own “anti-self” in the real world, and falls back into a dreamless sleep.

The final story in the collection, “Honey Pie,” is about a shy but increasingly successful short story writer named Junpei who has helped raise a female child named Sala with his longtime love from college, Sayoko. The problem is that Sayoko was married to their friend Takatsuki, a man she did not love, until Sala was two years old. This situation was partly a result of Junpei’s inability to articulate his feelings for Sayoko. “Honey Pie” begins with master storyteller, Junpei, telling Sala a story about a bear named Masakichi who made money selling honey, whom all the other bears hated. One particularly jealous bear was Tonkichi. In this case, Masakichi is the “seer” character, as he is Junpei’s guide to finding the courage to propose to Sayoko at the end of the story. The three of them go to a zoo the next day to see the bears, and Junpei finishes his story to Sala. Tonkichi trades salmon with Masakichi for honey, and they become friends, but Tonkichi gets captured and sold to the zoo. That night, Sayoko and Junpei sleep together, and Junpei decides to tell and write hopeful conclusions, with Tonkichi baking honey pies from the zoo to send to Masakichi so they can still be friends. Masakichi’s story prompts him to propose to Sayoko the next morning. The Kobe earthquake initiates this because Sala, after watching Kobe’s destruction on the news, begins to have nightmares about an “Earthquake Man” who comes with tiny boxes, waiting to squeeze them into the boxes. Metaphorically, of course, it represents social expectations and destiny, but here lies an opportunity to reject them. Junpei decides he is tougher than social conventions, that “no matter who it is, they won’t force [him] into meaningless boxes. Even if the sky falls down, even if the ground roars open” (237). The earthquake has strengthened him.

In comparison, Colorless Tsukuru Tazaki and His Years of Pilgrimage contains a longer journey, one centered on a personal disaster for the main character, Tsukuru. In high school, Tsukuru was part of an inseparable five-person group, two girls and three boys. All of them had names containing colors, except for Tsukuru, which means “to make”
in Japanese. Compared to his group of friends—which included a rugby player, a genius, a piano virtuoso, and a hilarious comedian—Tsukuru always felt uninteresting for his favorite pursuit, train station engineering. His friends stayed in their hometown, Nagoya, while Tsukuru went to Tokyo to study his passion. When he returns one college break, mysteriously, none of them answer his calls, and he realizes that they want nothing to do with him anymore. This hits him like a sarin gas terrorist attack, sending him into the depression about which the novel begins, “From July in his sophomore year to January of the following year, Tsukuru Tazaki lived thinking almost only about death”(3). Then the title reveals its meaning; Tsukuru worries if it was his “colorlessness” that caused his friends to abandon him.

With his health spiraling downward, Tsukuru descends into depression for five months, and comes out of it when he starts using the local gym and its pool. There, he begins to swim with an underclassman named Haida who inspires in Tsukuru a passion for classical music and philosophy. Haida’s name in Japanese contains the word “grey,” implying his wisdom for recognizing Tsukuru’s strengths, like when he notices that Tsukuru “likes to build things…like [his] name”(53). The long, 1500-meter slogs Haida leads Tsukuru through represent Tsukuru’s first journey in pulling strength from his psychological nightmare. After swimming, “again all the important muscles grew back, his spine stretched straighter, and even the color returned to his face…the boy once named Tsukuru Tazaki was dead”(60). Haida shows Tsukuru how to liberate his soul during one of their long midnight conversations. Haida says that he “wants to live a free life, the kind he can go where [he] wants when [he] wants, thinking only about the things [he] wants to”(63). Tsukuru admires Haida and this “freedom of thought.” However, Haida warns Tsukuru that to learn to think for himself “is as hard as dreaming intentionally. Normal people can’t pull it off”(63). In fact, this process is “the same as mastering a new language, putting its usage to heart.” Murakami furthers Haida’s position as a seer when Haida describes how prophets bring messages that people can use for “rebellion” against their thoughts and the “frames” or limitations of existence. The two then listen to Franz Liszt’s “Le Mal du Pays” (Homesickness), on a record named with the book’s title, “The Years
of Pilgrimage.” Haida describes it having “a sadness without reason,” a similar sensation to Tsukuru’s feelings when his friends abandoned him. In their friendship, Haida prepares Tsukuru for his journey to recover his friends by teaching him independence. He, too, disappears mysteriously one day, but this time, Tsukuru is stronger, and continues on his life for years before meeting his next “seer.”

Now cozy in his train station engineering job, Tsukuru drifts through relationships and life quietly. That is, until a woman he starts dating, Sara, gives him an ultimatum; he needs to connect with his four past friends, or he’ll never see her again. Again, she acts as a wise figure who sees Tsukuru’s buried pain and his past relationship troubles it caused. She knows that, to love another, one must love oneself. Sara hunts for the friends’ contact information on the Internet, and sends Tsukuru off on his pilgrimage to excavate his past. At that point, Tsukuru discovers a terrible hidden truth. One of his female friends was raped, and she had told the rest of their group that Tsukuru had done it. This is what prompted them to cut off all connections with him. The girl was found murdered in an apartment some years later.

Shocked, Tsukuru goes from Nagoya to Finland to reconcile with each of the remaining three friends. He learns that to his friends, he wasn’t the dull, colorless person he imagined himself to be, but rather, like the web of subways he watches and builds, he was their connector, the “ship’s anchor” that held the group down. He made them “feel safe… when [he] was with [them, they] felt their natural selves” (221). The girl living in Finland confesses she liked him because of this. One of the friends felt Tsukuru was the “handsome boy that all the parents loved.” Another friend reveals that after he left for Tokyo, the girl who was later killed may have been the first to sense their lives’ divergence and the group’s approaching collapse, and so, feeling trapped, she unconsciously blamed Tsukuru for her assault. Tsukuru realizes the effects he has had on his loved ones and his worth as a friend. But the reader also sees that this realization only comes through painful pilgrimages like these. Murakami contrasts Tsukuru with one of his friends, a former rugby player. The friend “was not used to confusion. His true worth showed on fixed fields, following fixed rules, playing with fixed teammates”(137). Tsukuru, on the other hand, has been dealing with
absurd puzzles thrown his way in life which has toughened him. His friend notices this, telling Tsukuru he has a “slim, masculine feel to [him. His] cheeks are hollow, and [his] eyes are deep and sharp.” In comparison, his once athletic friend complains of not being in shape enough now to comfortably run a mile. Tsukuru sees this and chooses not to talk about “those days that completely made new his mind and body… even if he told [his friend,] probably not even half of what he’d felt would get through”(138).

In the end, Tsukuru returns to Tokyo and reflects on his relationship to his family, his past, and his friends. He sits in Shinjuku Station, the busiest in the world, looks at the random movements of the crowd, and feels at peace. He is ready to love Sara fully.

Doubtless, as someone who can translate *The Great Gatsby* into a just-as-powerful Japanese version, Murakami loves language. As a researcher, as a short and long novelist, he knows concision in composition. However, he also adores the wispy, grey characters and words sliding around our vision. When Tsukuru talks with a woman in Finland, she says to him, “there are things in our lives too hard to explain, no matter what the language”(224). Because of the difficulty in expressing bizarre events in normal writing, his magical realism and random quests seek to copy the real world’s unreal happenings. Most tangibly, he uses natural disasters like earthquakes in *after the quake*. But he also explores “unreal” personal disasters. For us humans in an erratic universe, stories are what give us sense. ●
地面のとどろく国、火山の多い国、日本は昔から数々の災害をもって国家的帰属意識と結束を強めている国です。その例は古代から最近にまでおおよんでいます。1281年に台風が発生して、この最初のいわゆる「神風」は十四万人のモンゴルの兵隊を沈没させて、日本国を助けました。1995年のマグニチュード6.8の阪神大震災も日本史を変えました。日本はいろいろな災害により衝撃を受けました。一方では2011年の東日本大震災とそのあとの福島第一原子力発電所の事故、他方では1995年のオウム真理教により行われた地下鉄サリン事件のようなホームグロウン・テロリズム。たびたび「マジックリアルリズム」という文学形式もシュルレアルリズムも日常世界を混ぜて描く、村上春樹さんは日本でも世界中でもとても人気のある作家です。日本の精神がよく分かります。村上春樹さんの「神の子どもたちはみな踊る」という短編集と「色彩を持たない多崎つくる」という東日本大震災の後で書いた最新小説はマジックリアルリズムの程度と話の内容がさまざまですが、両作品は一つの考えを提案しています。それは災害が元気を取り戻せるものだということです。災害の突然もたらす喪失がありふれた考え方を震わせて歪める、そして我々の精神を鋭くします。それは自分たちを重ね上げ、自分たちのために代替現実と可能性を追いかける機会です。特にこの二冊の中では、村上さんのマジックリアルリズムが災害の不合理にぴったり合っていて、予期せぬ個人的な問題と環境災害は「予言者」という人物となって、主人公にもっと満足できる「巡礼」をする機会だと提示しています。

村上さんは「神の子どもたちはみな踊る」のあとがきの中で1995年の1月に阪神大震災と3月に地下鉄サリン事件からのインスピレーションを説明します。一つが天災でも、一つが人災でも、両方は結局日本をさざ波のショックさせた、と
彼は信じています。「我々はおおむね、自分の踏んでいる大地が揺らぎのないものだと信じている。あるいはいちいち信じるまでもなく、「自明の理」として受け入れている。しかし突然それは我々の足下で『液状化』してしまう… 目に見えない致死的な凶器が通勤する人々を無差別に襲う。」この短編集は村上さんの最初の三人称を使って書いた作品です。この一人称からの変化は、主人公の選択が個人的な旅の部分であるだけでなく、外的な力の行っている部分を強調します。すべての短編小説の中には阪神大震災が影響する主人公も主人公に自分の運命を支配できることにするという旅を差し出す「予言者」という人物もあります。こんな旅は、主人公の痛みを力に変えられると再考し、その力を持って満足感に人生を導きます。まず短編集の名前を付く短編小説から始めましょう。

ある朝、善也というパーティー好きな二十代の男が二日酔いで目覚ます。善也のお母さんは、高校時代によく「愛もなくまぐわって」男の人と付き合っていたあとで、耳たぶの欠けている医者と付き合っていました。そしてちゃんと避妊をしていたのに、善也を生みました。医者がお母さんを捨ててから、お母さんはある宗教の中で信仰を見つかりました。その信者たちは善也の処女懐胎が彼の神様の子供である象徴だと信じています。何も特に野球、うまくなかったと育ち行って、善也は神様が自分のお父さんだと信じませんでしたので、信仰を捨てました。阪神大震災のあとでお母さんはその地域に信者たちと一緒に援助の仕事に行きます。これを覚えて、善也は仕事をしないと決めて電車の中にいる耳たぶ欠けている昔から離れた生物学的な父親だと信じる男の人のあとを追い、水銀灯のある産業地帯を通して男の人を追い求めます。この水銀街灯は村上のマジックリアリズムとぴったりと会います：「あたりには人の生活の気配はなく、まるで夢の中でとりあげずしつらえられた架空の風景のようだ。」善也は「世界を威嚇するように厚く巡らされている」鉄条網に囲まれた巨大な「愛想のない」コンクリートの塀の
そばで追いつづけます。その場の中には車が堆積しており、男の人は暗い狭い路地に入り込み、「善也は少し躊躇したが、それでも男のあとを追って薄暗い闇の中に足を踏み入れた。」こんなのは「英雄の旅」ということではありませんか。向こう側にある穴から這って寂しい野球場に着くと、男が姿を消したと気付きます。この瞬間に我々は一般的にマジックリアリズムのような説明のつかない英雄の旅というところに見えることになったと理解します。ひとりぼっちの善也は人生の今までの関係はすべてがセクシーすぎるというとその付き合いが途中に踊ることが上手になるようになったということを思い出します。野球場で一人で踊りながら、世界と相互に連結される感じがして、自分が世界の善意に変わる可能性を悟ります。この理解はお母さんのセクシュアリティーとそこにあるより深い意味についての理解と類似点があります。お母さんは善也の出生でした。善也には、ナイトクラブの踊りです。両道は二人を神様に導きます。お母さんは信号を見つかりました。そして最後に善也は「神様」とつぶやきます。阪神大震災に母親との、父親のアイデンティティとの、生きがいとのつながりは揺れましたから、地震が起こらなかったならこのシーンの希望と善也の可能性のある救済も起こらなかったかもしれません。

***ここから読んでください：

「かえるくん、東京を救う」という短編小説の中で人間の大きさのカエルが巨大な地下に住んでいるみみずによる阪神大震災の同じような地震から東京を救うよう普通のローンコレクターの片桐の助けを求めます。このマジックリアリズムを強い効果を持たせるのはかえるくんの礼儀正しくて謙虚な話し方：「本来ならばアポイントメントをとってから来るべきところです。」珍しいですが、読者にも片桐にも魅力がありました。村上さんの主人公の例によって片桐は一匹狼で、中国人のマフィアやヤクザの間に単調なローンを集める仕事を
処理します。こんな人生はバブル後の日本の経済の、精神の憂鬱を表現します。しかし、「予言者」としてかえるくんが片桐をみみずと戦って人生の新しい意味に導きに来ます。戦いに会う場所に行く途中で、片桐は発射されて無意識状態になって、そのあと病院で起きてかえるくんが夢の地下の世界で行われた戦いの勝利を祝います。みみずに傷つけられた上に、「非自分」というような心の中にあるエントロピーを感じていていると告白して、少し眠って休みますが、病院の部屋を閉まれる虫の群れに食われます。片桐は目を覚ましてかえるくんの死で困っていますが、それから力を取って、今現実の世界で彼自身の「非自分」に対して戦えると悟って、夢を見なく眠り込みます。

最後の短編小説の「蜂蜜パイ」はおとなしくてゆっくりと成功を収めた「淳平」という短編小説作家についてです。淳平は女の子の沙羅を大学時代から長年に愛している友達と育っています。あいにく淳平が小夜子に愛を告白することができませんでしたから、小夜子は沙羅が二歳になるまで別の大学時代の友達の高槻と結婚していました。小説の初めにはマスター語り部の淳平は沙羅にまさきちという熊についての物語を語ってやります。まさきちは山の蜂蜜を売ってお金を稼ぎますけれども、ほかの熊たち特に「とんきち」はまさきちのことが嫌いです。今回、まさきちが淳平の小夜子に最後に告白できさせるガイドです。小夜子と沙羅と淳平は動物園に熊を見に行って、淳平は物語を語り続けてます。とんきちはまさきちと鮭を蜂蜜に交換して友達になるようになりますが、とんきちとは捕まえられて動物園に売られています。その夜、小夜子と淳平は一緒に寝てから淳平はもっと希望に満ちた終わりを書こうと決めます。とんきちは二匹の友情を守るように蜂蜜パイを焼いてまさきちに送ってあげます。まさきちの物語は淳平を小夜子に結婚を申し込ませます。阪神大震災とのつながりは、沙羅はテレビに神戸市の破壊を見たあとで悪夢を見るようになります。気持ち悪い「地震男」が彼らを狭い箱に閉じ込めようとするという夢なのです。もちろん象徴
的にこれは社会の義理や運命を表しますが、ここでそんなことを探す機会でもあります。淳平は自分がそれより強いと自分を拒む機会でもあります。淳平は自分がそれより強く、相手が誰であろうと、わけのわからない箱に入れさせたりはしない。たとえ空が落ちても、大地が音を立てて裂けても」と自分に言います。で、阪神大震災は淳平を強めました。

比べてみると、「色彩を持たない多崎つくると、彼の巡礼の年」の内容は主人公の「つくる」の個人的な人災と、そのあとのもっと長い旅です。高校時代につくるは仲良く離れられない五人組（2人の女、3人の男）の部分でした。皆は色つきの名前を持っていますが、つくるは持っていません。いつも駅の工学に対する興味で、様々な才能を持つ友達の陰にかくれて目立たないと感じていました。その四人は故郷の名古屋に残る時に、つくるは東京に大学で駅のことを勉強しに行きました。ある休み帰郷すると、説明もなく誰も電話を返事をしないし、四人はもう一度話したくないと言います。これはサリンガスのテロリストの攻撃のようにピシャリとひっぱたって、鬱病に落ち込みます。小説の初めはこう：「大学二年生の七月から、翌年の一月にかけて、多崎つくるはほとんど死ぬことにしか考えて生きていた。」それからタイトルが示す通り、つくるは自分の名前と人物の「色彩を持たないこと」がその四人がつくるを捨てた理由かどうかと悩みますが、まだ筋が通りません。

健康が急降下し、つくるは鬱病に落ち込みます。近くの体育館とそこであるプールに行き始めると治りました。そこで、年下の「灰田」と一緒に泳ぐようになります。灰田はクラシック音楽と哲学に対する興味をつくるの心に置きます。灰田の名前を象徴するのは知恵です。つくるの得意なところを認めます：「つくるさんは、何かを作るのが好きなんですね。名前通りに」もう一つの象徴、灰田のつくるの導く苦しい千百メートルの泳ぎは、つくるの第一の四人の友達によった精神的な悪夢から力を養う旅です。泳いでいるから、「再び必要な
筋肉がつき、背骨がまっすぐに伸び、顔にも血色が戻ってきた…多崎つくるという名のかつての少年は死んだ。二人の長い真夜中の会話の間に、つくるに精神を自由にし方を見せます：「好きなときに好きなところに行って、好きなだけ考えられるような、自由な生き方」つくるは灰田とこの「思考の自由」を感心します。しかし、灰田はつくるに「それは意図的に夢を見られるというのと同じくらい困難なことです。普通の人にはなかなかできません」と注意します。その上に、この行動は「新しい言語を習得し、その語法を暗記するのと同じような、自由な生き方」つくるは灰田とこの「思考の自由」を感心します。しかし、灰田はつくるに「それは意図的に夢を見られるというのと同じくらい困難なことです。普通の人にはなかなかできません」と注意します。その上に、この行動は「新しい言語を習得し、その語法を暗記するのと同じような、自由な生き方」つくるは灰田とこの「思考の自由」を感心します。しかし、灰田はつくるに「それは意図的に夢を見られるというのと同じくらい困難なことです。普通の人にはなかなかできません」と注意します。その上に、この行動は「新しい言語を習得し、その語法を暗記するのと同じような、自由な生き方」つくるは灰田とこの「思考の自由」を感心します。しかし、灰田はつくるに「それは意図的に夢を見られるというのと同じくらい困難なことです。普通の人にはなかなかできません」と注意します。その上に、この行動は「新しい言語を習得し、その語法を暗記するのと同じような、自由な生き方」つくるは灰田とこの「思考の自由」を感心します。しかし、灰田はつくるに「それは意図的に夢を見られるというのと同じくらい困難なことです。普通の人にはなかなかできません」と注意します。その上に、この行動は「新しい言語を習得し、その語法を暗記するのと同じような、自由な生き方」つくるは灰田とこの「思考の自由」を感心します。しかし、灰田はつくるに「それは意図的に夢を見られるというのと同じくらい困難なことです。普通の人にはなかなかできません」と注意します。その上に、この行動は「新しい言語を習得し、その語法を暗記するのと同じような、自由な生き方」つくるは灰田とこの「思考の自由」を感心します。しかし、灰田はつくるに「それは意図的に夢を見られるというのと同じくらい困難なことです。普通の人にはなかなかできません」と注意します。その上に、この行動は「新しい言語を習得し、その語法を暗記するのと同じような、自由な生き方」つくるは灰田とこの「思考の自由」を感心します。しかし、灰田はつくるに「それは意図的に夢を見られるというのと同じくらい困難なことです。普通の人にはなかなかできません」と注意します。その上に、この行動は「新しい言語を習得し、その語法を暗記するのと同じような、自由な生き方」つくるは灰田とこの「思考の自由」を感心します。しかし、灰田はつくるに「それは意図的に夢を見られるというのと同じくらい困難なことです。普通の人にはなかなかできません」と注意します。その上に、この行動は「新しい言語を習得し、その語法を暗記するのと同じような、自由な生き方」つくるは灰田とこの「思考の自由」を感心します。しかし、灰田はつくるに「それは意図的に夢を見られるというのと同じくらい困難なことです。普通の人にはなかなかできません」と注意します。その上に、この行動は「新しい言語を習得し、その語法を暗記するのと同じような、自由な生き方」つくるは灰田とこの「思考の自由」を感心します。しかし、灰田はつくるに「それは意図的に夢を見られるというのと同じくらい困難なことです。普通の人にはなかなかできません」と注意します。その上に、この行動は「新しい言語を習得し、その語法を暗記するのと同じような、自由な生き方」つくるは灰田とこの「思考の自由」を感心します。しかし、灰田はつくるに「それは意図的に夢を見られるというのと同じくらい困難なことです。普通の人にはなかなかできません」と注意します。その上に、この行動は「新しい言語を習得し、その語法を暗記するのと同じような、自由な生き方」つくるは灰田とこの「思考の自由」を感心します。しかし、灰田はつくるに「それは意図的に夢を見られるというのと同じくらい困難なことです。普通の人にはなかなかできません」と注意します。その上に、この行動は「新しい言語を習得し、その語法を暗記するのと同じような、自由な生き方」つくるは灰田とこの「思考の自由」を感心します。しかし、灰田はつくるに「それは意図的に夢を見られるというのと同じくらい困難なことです。普通の人にはなかなかできません」と注意します。その上に、この行動は「新しい言語を習得し、その語法を暗記するのと同じような、自由な生き方」つくるは灰田とこの「思考の自由」を感心します。しかし、灰田はつくるに「それは意図的に夢を見られるというのと同じくらい困難なことです。普通の人にはなかなかできません」と注意します。その上に、この行動は「新しい言語を習得し、その語法を暗記するのと同じような、自由な生き方」つくるは灰田とこの「思考の自由」を感心します。しかし、灰田はつくるに「それは意図的に夢を見られるというのと同じくらい困難なことです。普通の人にはなかなかできません」と注意します。その上に、この行動は「新しい言語を習得し、その語法を暗記するのと同じような、自由な生き方」つくるは灰田とこの「思考の自由」を感心します。しかし、灰田はつくるに「それは意図的に夢を見られるというのと同じくらい困難なことです。普通の人にはなかなかできません」と注意します。その上に、この行動は「新しい言語を習得し、その語法を暗記するのと同じような、自由な生き方」つくるは灰田とこの「思考の自由」を感心します。しかし、灰田はつくるに「それは意図的に夢を見られるというのと同じくらい困難なことです。普通の人にはなかなかできません」と注意します。その上に、この行動は「新しい言語を習得し、その語法を暗記するのと同じような、自由な生き方」つくるは灰田とこの「思考の自由」を感心します。しかし、灰田はつくるに「それは意図的に夢を見られるというのと同じくらい困難なことです。普通の人にはなかなかできません」と注意します。その上に、この行動は「新しい言語を習得し、その語法を暗記するのと同じような、自由な生き方」つくるは灰田とこの「思考の自由」を感心します。しかし、灰田はつくるに「それは意図的に夢を見られるというのと同じくらい困難なことです。普通の人にはなかなかできません」と注意します。その上に、この行動は「新しい言語を習得し、その語法を暗記するのと同じような、自由な生き方」つくるは灰田とこの「思考の自由」を感心します。しかし、灰田はつくるに「それは意図的に夢を見られるというのと同じくらい困難なことです。普通の人にはなかなかできません」と注意します。その上に、この行動は「新しい言語を習得し、その語法を暗記するのと同じような、自由な生き方」つくるは灰田とこの「思考の自由」を感心します。しかし、灰田はつくるに「それは意図的に夢を見られるというのと同じくらい困難なことです。普通の人にはなかなかできません」と注意します。その上に、この行動は「新しい言語を習得し、その語法を暗記するのと同じような、自由な生き方」つくるは灰田とこの「思考の自由」を感心します。しかし、灰田はつくるに「それは意図的に夢を見られるというのと同じくらい困難なことです。普通の人にはなかなかできません」と注意します。その上に、この行動は「新しい言語を習得し、その語法を暗記するのと同じような、自由な生き方」つくるは灰田とこの「思考の自由」を感心します。しかし、灰田はつくるに「それは意図的に夢を見られるというのと同じくらい困難なことです。普通の人にはなかなかできません」と注意します。その上に、この行動は「新しい言語を習得し、その語法を暗記するのと同じような、自由な生き方」つくるは灰田とこの「思考の自由」を感心します。しかし、灰田はつくるに「それは意図的に夢を見られるというのと同じくらい困難なことです。普通の人にはなかなかできません」と注意します。その上に、この行動は「新しい言語を習得し、その語法を暗記するのと同じような、自由な生き方」つくるは灰田とこの「思考の自由」を感心します。しかし、灰田はつくるに「それは意図的に夢を見られるというのと同じくらい困難なことです。普通の人にはなかなかできません」と注意します。その上に、この行動は「新しい言語を習得し、その語法を暗記するのと同じような、自由な生き方」つくるは灰田とこの「思考の自由」を感心します。しかし、灰田はつくるに「それは意図的に夢を見られるというのと同じくらい困難なことです。普通の人にはなかなかできません」と注意します。その上に、この行動は「新しい言語を習得し、その語法を暗記するのと同じような、自由な生き方」つくるは灰田とこの「思考の自由」を感心します。しかし、灰田はつくるに「それは意図的に夢を見られるというのと同じくらい困難なことです。普通の人にはなかなかできません」と注意します。その上に、この行動は「新しい言語を習得し、その語法を暗記するのと同じような、自由な生き方」つくるは灰田とこの「思考の自由」を感心します。しかし、灰田はつくるに「それは意図的に夢を見られるというのと同じくらい困難なことです。普通の人にはなかなかできません」と注意します。その上に、この行動は「新しい言語を習得し、その語法を暗記するのと同じような、自由な生き方」つくるは灰田とこの「思考の自由」を感心します。しかし、灰田はつくるに「それは意図的に夢を見られるというのと同じくらい困難なことです。普通の人にはなかなかできません」と注意します。その上に、この行動は「新しい言語を習得し、その語法を暗記するのと同じような、自由な生き方」つくるは灰田とこの「思考の自由」を感心します。しかし、灰田はつくるに「それは意図的に夢を見られるというのと同じくらい困難なことです。普通の人にはなかなかできません」と注意します。その上に、この行動は「新しい言語を習得し、その語法を暗記するのと同じような、自由な生き方」つくるは灰田とこの「思考の自由」を感心します。しかし、灰田はつくるに「それは意図的に夢を見られるというのと同じくらい困難なことです。普通の人にはなかなかできません」と注意します。その上に、この行動は「新しい言語を習得し、その語法を暗記するのと同じような、自由な生き方」つくるは灰田とこの「思考の自由」を感心します。しかし、灰田はつくるに「それは意図的に夢を見られるというのと同じくらい困難なことです。普通の人にはなかなかできません」と注意します。その上に、この行動は「新しい言語を習得し、その語法を暗記するのと同じような、自由な生き方」つくるは灰田とこの「思考の自由」を感心します。しかし、灰田はつくるに「それは意図的に夢を見られるいう
驚いてつくるは残りの三人と仲直りしに名古屋からフィンランドまで行きます。もう一つの驚くのは、友達にとって、つくるは自分を想像した、ぼんやりと色彩を持たない人ではないし、つくるの作る地下鉄の網のように、「船の錨」のように、彼は組のつながる力です。「みんなの心を落ち着かせていく... おまえがそこにいるだけで、おれたちはうまく自然におれたちでいられるようなところがあったんだ。」フィンランドに住んでいる女がこのつくるの特徴で彼のことが好きだったと明らかにします。友達の一人はつくるは「育ち良いハンサムボーイ」だったし、一番いい聞き手だったと思います。もう一人から聞く発見は、つくるが東京に出発したあとで、殺された女が最初に組員たち的人生の近づいていた不調和、そして組の発散に気付いたかもしれませんので、罠にかかっているような感じで、無意識的にレイプの責任をつくるのせいにしてしまいました。こんな発見のあとで、つくるは自分の影響と友人としての価値を悟ります。しかしこの悟りはこんな苦しい巡礼のあとしか来ないと見えます。村上さんはつくるとかつてのラグビーしていた友達を比較します。友達は「乱されるとそれに馴れていない。定められたフィールドの中で、定められたルールに従って、定められたメンバーと行動するときに、彼の真価が最も良発揮される。」逆に、つくるは●
Conceptions of Conception: Definitions of the Beginning of Life and Their Effect on Abortion Regulation

Caroline Wheeler

“One's philosophy, one's experiences, one's religious training, one's attitude toward life... are all likely to influence and to color one's thinking and conclusions about abortion.”

-Judge Harold A. Blackmun, US Supreme Court, 1973

In August 1969, a Texas woman named Norma McCorvey discovered she was pregnant. Separated from her husband, whom she had married when she was only sixteen years old, McCorvey was in no financial state to raise this child. She knew that the only way she could prevent the birth of this baby, which would be her third, was to get an abortion. However, at the time, Texas law prohibited abortion except in the case of rape and incest. In her desperation and at the encouragement of friends, McCorvey decided to lie about her pregnancy, saying that she had been raped while walking home from work late one night.

At the same time, two lawyers named Linda Coffee and Sarah Weddington were searching for a pregnant woman to help them bring a case against the State of Texas regarding its abortion policy. McCorvey, who had been unable to receive sympathy from her doctor, contacted Coffee and Weddington, who were willing to make that case. In order to protect McCorvey’s identity, the three women came up with an alias: Jane Roe.

The following summer, in June 1970, the Texas State abortion law was ruled unconstitutional. The court stated that “the privacy of a woman should be protected; she should be able to decide what to do with her

This paper was written for Peter Brown’s Global Scholars: Population & Progress class in the spring of 2014.
body.” However, the court did not issue an injunction, in this case an order requiring the defendant to act on the court’s ruling. Therefore, the State of Texas was still able to enforce its antiabortion law. Coffee, Weddington, and Roe (McCorvey) then appealed to the United States Supreme Court in a case known today as *Roe v. Wade*. Several years later, in 1973, the Supreme Court ruled that women have a qualified right to get an abortion. *Roe v. Wade* is one of the most significant and controversial Supreme Court cases in American history, and its outcome remains a momentous milestone in the ongoing abortion debate in the United States.

The heated debate over abortion that has raged in the United States is not unique. All over the world, governments, NGO’s, religious groups, human rights workers, and various political groups alike are working to shape regulations on abortion. While many policies are similar, the motives behind these regulations are quite varied. What could account for this variety? The most common argument against abortion is that the act is murder. However, contrary to the widespread prohibitions against murder across different religions and political systems worldwide, the contrasting scientific, religious, and political viewpoints regarding the beginning of life have led to a lack of consensus concerning abortion policies. This paper will examine the different definitions of when life begins offered by embryology, indigenous traditions, and Abrahamic and Buddhist religions. Then it will explore how these definitions affect political regulations on abortion in a variety of countries, used as case studies, throughout the world. Through comparison and analysis of the dominant religious nature and abortion policy in these countries, this paper will demonstrate that restrictions and allowances for abortion correlate with the religiosity of a country.

**Historical Overview**

First, from time immemorial, people have wondered the question, “what defines a human self?” Martin J. Buss, professor emeritus of Religion at Emory University, states in his article, *The Beginning of Human Life as an Ethical Problem*, that “in order to determine when his life begins, it is necessary to ask what man [himself] is . . . A useful
way of understanding man . . . [is to see him] as one who refers back to himself. This selfhood is carried by culture, which is a system inherited not biologically, but by direct transmission from one person to another.”¹

Throughout history, many different cultures and peoples have tackled the question of when life begins in different ways. Without the technology and scientific knowledge we now possess, the beliefs and rituals of indigenous cultures surrounding childbirth were quite varied. For example, the Navajo people, a Native American tribe of the southwestern United States, did not consider a child to be alive until it had laughed for the first time. Prior to its first laugh, the baby was kept in a cradleboard, which simulated pregnancy. Another example is some tribes in New Guinea who believed that all children were female until they reached puberty. At that milestone, the male tribe members would take the boys away and have a birthing ceremony. This marked the beginning of their lives.²

Additionally, cultures around the world often treated individuals as extensions of their families or tribes rather than as people with specific individual rights. The Romans, for example, didn’t consider a baby as an individual, so children born with deformities, or children without accepted parentage were left in the wilderness to perish. This policy extended even into modern times for children without official parentage; the children of unwed mothers, for example, were never accorded full status in a society even if they managed to survive infancy.

The previous examples have focused on the official birth of the child. However, in some cultures, the act of giving birth marked the beginning of life. For example, in ancient China, a woman had two lives, the first of which began when she was born, the second when she gave birth to her first son. At this point, she was given a new name and was said to have been born again.³

A particularly intriguing example of indigenous views on conception and the beginning of life comes from the Aboriginal tribes of Australia. They believed that a “spirit child,” the Western equivalent
of an individual’s soul, chooses his or her parents. The spirit child was said to enter the chosen woman as a “complete entity who…originated at some time in the long distant past. The new life is immeasurably more ancient and completely independent of any living person.” Even after learning about the biological process of sexual intercourse and embryonic development, Aboriginals continue to believe that, while biological conception may “prepare the way for the spirit child’s entry into the womb,” the combination of sperm and egg is not the only cause of conception. They believe that the presence of a spirit child and his or her role in selecting parents is necessary in order to allow biology to take its course. For the Aboriginals, the question of when life begins anticipates the nature of debate in the West and elsewhere: the question of when life begins is not primarily biological but spiritual.

Biological Beginning of Life

Over time, however, as the debate surrounding abortion has unfolded, there has been an increasing interest in the science surrounding the beginning of life. Lawyers, politicians, biologists, women all ask the same question, when does a new life begin in biological terms? By turning to science, these people hope to receive a concrete definition of the beginning of human life that will help them quickly solve the controversy over abortion once and for all. Unfortunately, however, biology yields an ambiguous answer to the question of when life actually begins. While biologists are in agreement concerning the stages of fetal development, they do not necessarily agree at what stage a new life begins, as the very definition of “life” is as much a cultural as scientific construct.

Despite this disagreement, the different stages of embryo development have been studied and accepted as scientific fact, worldwide. During the first stage, fertilization, the parental gametes, sperm and egg, or more specifically their nuclei, join together to form a zygote, a single cell organism with a unique genetic makeup.

After the zygote has been formed, the embryo enters the cleavage stage. During this time, the embryo undergoes several important divisions.
First the zygote divides symmetrically in half forming embryo in the two cell stage, then divides in half again forming the four cell, and then the eight cell stage. Once it has reached the eight cell stage, the embryo ceases to divide symmetrically, and instead begins to divide randomly forming a blastocyst, or hollow sphere of cells, called blastomeres, which surround a fluid filled center, called the blastocoel.

Cleavage is a particularly important stage of development. As Senior Research Associate and Professor Emeritus of Developmental Biology at Swarthmore College Scott F. Gilbert states in his book Developmental Biology:

*The cytoplasm of the oocyte (egg) frequently contains morphogenic determinants that become segregated into specific cells during cleavage. These determinants ultimately lead to the activation or repression of specific genes and thereby confer certain properties to the cells that incorporate them.*

Thus, in order for an embryo to develop properly, these determinant factors must be divided correctly. This is demonstrated in Hans Spemann’s experiment in which he observed the effects of restricting developmental potential. In his experiment, Spemann first observed the normal cleavage pattern of a frog embryo. He found that when the embryo divided naturally and the animal pole, vegetal pole, and grey crescent were split evenly, the two daughter cells developed normally. However, when he used string to induce an unnatural division with the entirety of the grey crescent in one half, he found that only the daughter cell containing the grey crescent developed normally, while the other did not. Based on the results of his experiment, Spemann concluded that embryonic fates are affected early on in development by the distribution of determinants and the pattern of cleavage. The cleavage stage is also when the embryo implants in the wall of its mother’s uterus.

Once the blastula is fully formed and implantation has occurred, the embryo enters the gastrulation stage of development. During gastrulation, the blastula rearranges itself into a gastrula, or structure
comprised of three germ layers: ectoderm, mesoderm, and endoderm. Each of these germ layers will eventually bring about specific tissues and organs. For example, the ectoderm will create the epidermis (outer layer of skin) and other tissues central to the nervous system. The mesoderm will ultimately become blood, bone, muscle, and connective tissues (parts of the muscular, skeletal, and circulatory systems). Finally, the endoderm will become the organs and lining of the digestive system and respiratory system.

Following gastrulation, the embryo goes through neurulation, the formation of the neural tube, which eventually develops into key parts of the central nervous system: the spinal cord and brain. Then, finally, the embryo enters the organogenesis stage, during which specific organs and other structures are formed. By the end of this stage, the embryo becomes a fully functional organism capable of independent survival.

Despite the widespread acceptance of the findings of embryology throughout the scientific community, biologists disagree on precisely when, in the midst of the many stages of development, a new life begins. Of course, the answer to this question depends on how one defines human life. For example, if one believed that life begins when the fetal brain develops, then that person would claim that a fetus is not alive until the two months after conception, when the telecephalon forms a “conspicuous [slightly bilobed] bulge dorsally in front of the eye rudiment . . . [which is] the first indication of the future hemispheres of the brain.”

Some biologists believe that a new life begins at the moment of conception. In her article *When Does Human Life Begin? A Scientific Perspective*, Maureen Condic, an associate professor of Neurobiology and Anatomy at the University of Utah School of Medicine, argues that a new life is formed mere seconds after the binding of a sperm and egg, which creates a zygote, or single cell embryo, genetically unique from either parent gamete. Following the membrane fusion of sperm and egg, the zygote undergoes changes in its ionic composition that alter the chemical makeup of the zona pellicida, an “acellular structure
surrounding the zygote,” which serves to block other spermatozoa from entering the already fertilized egg. The behavior as well as the unique genetic makeup of the zygote are so drastically different from that of either parental gamete, Condic argues, the zygote appears to be its own individual organism.9

Furthermore, from the beginning stages of embryotic development, the zygote displays a pattern of “organization towards that of the production of a mature human body.”10 She explains that an organism is defined as two things: “a complex structure of interdependent . . . elements whose relations . . . are largely determined by their function” and “an individual constituted to carry out the activities of life by means of organs separate in function but mutually dependent.”11 From the early stages of development, embryonic cells are divided into three classes: ectoderm, mesoderm, and endoderm. Each of these categories will give rise to different aspects of the embryo's ultimate body. Thus, Condic argues, the embryo displays a human pattern of developmental behavior from the moment of conception.

Condic is not alone in her belief that life begins at conception. Dr. Alfred Bongioanni, a professor of pediatrics and obstetrics at the University of Pennsylvania stated during a United States Senate Judiciary Subcommittee hearing: “I submit that human life is present throughout [the] entire sequence from conception to adulthood and that any interruption at any point throughout this time constitutes a termination of human life.”12

However, the idea that life begins at conception is by no means a universally accepted fact. For other biologists, it is impossible to define when a new human life begins due to the cyclical nature of reproduction. These biologists point out that “the lifecycle of organisms that produce asexually does not involve conception,” so therefore the belief that life begins at conception is moot.13 Rather, as Robert Wyman points out in a lecture on biology and the history of abortion given at Yale University, life is a cycle and “fertilization is one event of that cycle.” Wyman states that the closest one can get to pinpointing when a person’s life begins is
inside the womb of that individual’s maternal grandmother, for that is when one’s eventual birth becomes a possibility. However, if one were to ask Wyman when life really begins, he would answer approximately 3.5 billion years ago, when the first unicellular organisms appeared on Earth.

For yet other biologists, such as Clifford Grobstein, a professor of Science, Technology, and Public Affairs at the University of California San Diego, the more relevant question is not when exactly life begins, but rather at what stage of development will an embryo that is “destined to acquire . . . attributes of a human being” be provided the same rights and protections given to those who have already been born. In answer to this question, Grobstein examines two commonly, though not unanimously, accepted aspects that help define fully-fledged human beings. The first is “wholeness in the sense of indivisibility.” In reaction to this characteristic, Grobstein argues that “scientifically, this attribute is not present in the mammalian zygote.” Therefore, by Grobstein’s definition, Condic’s and Bongionni’s zygote is not a “life” at all. The second attribute Grobstein examines is sentience, or the capacity to “perceive subjectivity.” Grobstein states that sentience is not present until the central nervous system develops (around the fifth week), therefore human life does not begin until sometime after conception.

Grobstein’s conclusion relates to the argument from the Roe v. Wade case: abortion is considered legal up to the point at which the fetus is able to survive independent of its host (the mother), after which the government has an interest in preserving its life.

In conclusion, the close study of the beginnings of life demonstrates that emerging life is a series of potentials, each made possible by previous potentials being fulfilled. Biologists disagree on which of the many steps in this process should be deemed the exact point at which life begins. Thus, science does not yield a single, conclusive answer to truly inform the cultural and legal debates about abortion.

Conception, Sin, and Abortion in the Abrahamic Traditions

Modern states—regardless of whether they are democratic in nature—have to deal with traditions that underpin their societies and national identities that may work to limit or constrict the rights of certain groups
within their societies. In the case of abortion, the issue of individual rights to bodily integrity and autonomy bump up against traditional concepts of the role of women as primarily wives and mothers. In the issue of abortion, governments have to navigate the concept of competing rights. Often, governments turn to the language and written traditions of the religion that defines or dominates in their societies. This section will focus on four dominant religions: the three Abrahamic traditions from the West and one from the east, Buddhism. Combined, these four religions affect the daily lives of a vast percentage of the world’s population. The three Abrahamic religions—Judaism, Christianity, and Islam—all stem from the Middle East and are named for a shared mythic genealogy. In addition to their common heritage, these three religions have several other factors in common, such as a monotheistic nature, in which a single God is worshipped, and ethical orientation, or an inherent belief in good versus evil. Another shared characteristic, one that will become significant later on in his paper, is the fact that all three Abrahamic religions are based upon central scriptures which are believed to reveal the words and intentions of God.

**Judaism**

In Judaism, the oldest Abrahamic religion, a fetus is not equivalent to an infant—an important difference when it comes to abortion policy. The famous Talmudist Rashi once said “whatever has not come forth into the light of the world is not a full human life.” In Judaism, the fetus is seen as potential life, rather than its own individual. Just as Jewish views on conception differ from those of Christianity, so do Jewish ideas on sin. Christians believe that all people come into the world burdened with original sin, that they are born guilty. Jews, on the other hand, believe that people are born innocent and that sin occurs as a result of individuals’ actions. Indeed, for both religions some acts, such as murder and stealing, are considered inherently sinful. However, as Tomas Silber points out in his article *Abortion: A Jewish View*, “the law of homicide in the Torah refers to . . . any living man. This excludes the fetus in the womb which is [not considered a person] until it is born.”
Under Jewish law, a fetus is seen as a “part of its mother rather than an independent entity.” Therefore, rather than viewing the issue in relation to the fetus, abortion is viewed in the context of the mother. For example, in Exodus:

*If men were fighting and someone struck a pregnant woman and she miscarried but she herself lived, he will be subject to a fine, as much as the woman’s husband shall request and as the judges decree. If, however, her death shall follow, let him pay a soul for a soul, an eye for an eye, a tooth for a tooth, a hand for a hand, a foot for a foot, a burning for a burning, a wound for a wound, a bruise for a bruise.*

Similarly, Judaism forbids self-mutilation and suicide, two acts that can be applied to the act of abortion. However, as Silber argues, “it is acceptable to have a limb amputated for the sake of the whole individual, [a] concept [that] is applied to the [legal] permission for abortion.”

**Christianity**

For a long time, many Christians have believed that life begins at conception. This idea is true of the Catholic Church and many other denominations worldwide. However, nowhere in its scriptures does Christianity directly define when life begins. Indeed, early Christians, influenced by the ideas of Plato, believed in the theory that the human soul does not enter the body until birth. Aristotle, a student of Plato’s, thought that the “time of entrance of a human soul was . . . forty days after conception for a male and ninety days for a female.”

One topic that is addressed in the Bible, however, is sin. According to the scriptures, there are many ways to commit sin in Christianity, but one of the most serious sins is murder. Since murder can be defined as destroying life, it is clear to see where the idea that abortion is a sin originated in the Christian tradition. In her article *Abortion in Italy*, Lesley Caldwell states that “abortion [involves] the destruction of human life and [therefore] can never be accepted regardless of the difficulties this position imposes [meaning] that the church could not admit the possibility of legal abortion.”
The idea of abortion within the Christian faith, particularly in Catholicism, has been considered a sin and a blatant “refusal of a divine gift.” Even before there was technology that could give potential evidence that life begins at conception, Christians believed that abortion constituted “proleptic murder”—the destruction of a potential life. Indeed, this belief stems back to Tertullian, a Christian author born in 160 CE, who proposed the concept of *homo est qui est futurus*, or “the being that will be a human is already to be regarded as human.”

Much more recently, in 1988, the Vatican released an official ruling that defined abortion as “any method used to terminate a pregnancy from the moment of conception,” thereby making it, as well as most contraceptives, illegal according to the Catholic Church. It is important to recognize that in predominantly Catholic countries, and in Italy in particular, the official opinions of the Vatican have powerful influence over both social and political aspects of daily life.

In the same 1988 statement, the Vatican also ruled that the single situation in which abortion would not call for excommunication was if the mother “urgently [required] a life-or-death procedure with the unwanted consequence of ending her pregnancy.” John Allen described such a situation as the “double effect” in which the positive, intended outcome (saving the mother) results in the negative, unintended outcome (death of the fetus).

However, in some Catholic countries, such as El Salvador, laws regarding abortion are so restrictive that most doctors refuse to abort even to protect the life of the mother, as they fear state prosecution more than they desire to preserve the life of their female patients. More recently, even in the United States, a hospital in Texas kept a brain-dead woman alive because it feared that, under Texas state law, they would be criminally liable if they did not do everything possible to preserve the life of the brain-dead woman’s fetus. Only after the woman’s family sued, and outside experts indicated that the fetus was not developing normally, did the hospital relent.
Islam

The third Abrahamic religion, Islam, is predominant throughout the Middle East, North Africa and much of Southeast Asia. A significant aspect of Islamic principle is that of individual interpretation. Leila Hessini, a global feminist leader and activist, states that “Muslims are encouraged to read and analyze traditional religious sources to find solutions to contemporary problems.”

Muslims believe that “ensoulment,” the moment at which an individual’s soul enters his or her body, marks the beginning of a new life. This belief stems from a particular verse in the Koran, the central Islamic scripture, which discusses embryonic development:

*Man We did create from a quintessence (of clay); then We placed him as (a drop of) sperm in a place of rest, firmly fixed; then We made the sperm into a clot of congealed blood; then of that clot We made a (fetus) lump; then We made out of that lump bones and clothed the bones with flesh; then We developed out of it another creature. So blessed be Allah the Best to create!*

The majority of Islamic schools of thought believe that ensoulment of the fetus occurs “40, 90, or 120 days after conception.” Interestingly, the Muslim concept of ensoulment is believed to occur approximately at the same time when the central nervous system (spinal cord and brain) develops in the fetus, a stage which begins at around five weeks after conception.

Similar to Christianity and Judaism, Islam also regards murder as a sin. The Koran explicitly states: “do not slay the soul which Allah has forbidden except in the requirements of justice.”

Concerning the issue of abortion, Muslim cultures generally forbid it after the fetus has achieved ensoulment, since the act would at that time constitute murder. The only exception to this rule is when abortion is required to save the woman’s life. As Hessini states in her article *Abortion and Islam: Policies and Practice in the Middle East*
and North Africa, “it is accepted [throughout Muslim countries] that maternal life takes precedence [over the life of the fetus, even one that has reached ensoulment], at least until the fetus achieves the status of person.” On the other hand, abortion is strictly forbidden without the consent of the father and in cases when the pregnancy is a result of illicit sexual activity. In essence, Muslim views on abortion are much more similar to those of Judaism, in which the life of the mother is considered more valuable than the life of the fetus (which, is not even considered “alive” until sometime after conception), than the beliefs of the Catholic Church.

Conception, Sin, and Abortion in Buddhism

Unlike the three main Western religions, Christianity, Judaism, and Islam, Buddhism is not an Abrahamic tradition. Rather than worshipping a single, omnipotent God, Buddhism is a nontheistic religion based on the ideas of a man named Siddhartha Gautama, or simply Buddha, the Pali word for one who is “awake.”

Buddhism revolves around the idea of suffering. The goal of Buddhism is to lessen individual suffering, increase happiness and, ultimately, reach a state of nirvana. Therefore, sin in Buddhism is considered to be anything that causes harm or suffering. A key aspect of Buddhism is the concept of karma, the belief that intentions and actions of individual’s will affect that being’s future.

Buddhists do not believe that life begins at conception like Christians, nor do they think it begins at birth like Jews. Instead, Buddhism states that a new human life begins when “citta (mind or thought) arises, when the first consciousness manifests [itself],” when it is able to feel suffering. For Buddhists, it would be false to state that this consciousness “descends into the mother’s womb at the very moment of parental union.” Instead, according to the teachings of Buddha, “[the embryo only] takes place through the union of three things—the union of the mother and father, [when] the mother is in season, and the gandhabba (stream of consciousness) is present.” The Buddhist concept of consciousness can be likened to Grabstein’s idea of sentience—a new life begins when the fetus is able to feel and react to stimuli.
When it comes to Buddhist sin, the concept of karma plays a significant role. Buddhists tend to look at life as a balance between good and bad actions, which in turn affects their “karmic status.” Therefore, one’s actions cannot be examined on “an absolute scale of good and evil, but rather . . . understood in terms of a relative scale.” Consequently, for Buddhists, the degree of sin, which can be influenced by the sinner’s motivation or reasoning, is an important factor in determining how a particular action will affect their karmic status.

According to Brahm, “only when the embryo-fetus first shows sensitivity to pleasure and pain and first shows will (such as by a purposeful shrinking away from a painful stimulus) has consciousness . . . first manifested and the new human life started.” Before the citta first develops in the fetus, abortion is not technically considered a sin, although it is by no means seen as desirable. However, once the stream of consciousness is established in a fetus, abortion is viewed as a “life-destroying act . . . [constituting] a serious sin.” However, Buddhists believe that there are “different levels of sin depending on the circumstances and intentions with which the act is performed.” Therefore, it is understood that when a Buddhist woman gets an abortion, she recognizes that her actions are considered harmful, but is nevertheless willing to accept the karmic consequences of her actions.

Abortion Policies by Country

So, how do these religious traditions affect national policies? In his paper Through Tinted Glasses: Religion, Worldviews, and Abortion Attitudes, Dr. Michael O. Emerson, a professor of Sociology at Rice University, states that “religion provides answers to questions of ultimate meaning. It also provides guidelines for how to live in the present world. By outlining what ought to be and by creating and reinforcing group norms through interaction, religion has a substantial influence on personal worldviews.” Although religion is by no means the only factor considered when coming up with national abortion policies, there is an undeniable correlation between the religiosity—“the intensity and consistency of a persons’ practice of their religion”—of
a population and that country’s regulations on abortion. This section will examine the abortion policies in eight different countries from all around the world. These countries, which are used as case studies, include Australia, Finland, Iran, Ireland, Israel, Japan, Thailand, and the United States.

Throughout the world, there are three general “levels” of abortion policies. The first and most common is very restrictive. In countries where the abortion regulations are deemed very restrictive, abortion is either prohibited completely or allowed only to save the mother’s life. The Republic of Ireland is a good example of a very restrictive country.

The next level is limited. Limited policies, like those in Japan, Iran, and Thailand, disapprove of abortion on principle, yet make a point of allowing it in order to preserve the mother’s health. Since the definition of “preserving health” is a very broad, this particular category can be applied to many policies of differing restrictiveness.

The third and final level is conditional. Countries with conditional abortion policies consider several factors before deciding whether or not a woman will be allowed to get an abortion. Such factors include, but are not limited to, preserving the life and health of the mother, the circumstances of the pregnancy (rape, incest, etc.), and various socio-economic factors. Two countries with conditional abortion policies are Israel and Finland.

Republic of Ireland

In 2013, the Irish parliament passed a new law that allowed abortion under circumstances when the pregnancy endangers the life of the woman. According to the regulation, a pregnant woman’s life could be endangered in two ways. The first is if she suffered from a severe physical illness. In this case, the woman would have to have the approval of two physicians before she could qualify for an abortion. One of these physicians must be an obstetrician and the other must be a specialist in the field relevant to her illness. For example, if a woman suffered from cancer, she would have to get the approval of her obstetrician and an
oncologist. However, if the illness caused the woman to be in a medical emergency, the approval of only one physician is required in order to save her life.\textsuperscript{37}

The second situation in which a woman’s life is considered endangered is if her pregnancy causes her to contemplate committing suicide. In cases such as these, the approval of three physicians—one obstetrician and two psychiatrists, one with experience treating women during or after pregnancy—must be obtained before she can qualify for abortion.\textsuperscript{38}

The very restrictive abortion policy in the Irish Republic is heavily influenced by Roman Catholicism. Since, for Catholics, it is the soul which “distinguishes human beings from other entities and the soul enters the fetus at conception,” harming the fetus in any way is considered a sin, and abortion is considered to be on par with murder. Additionally, there is an expectation in Catholicism that women are expected to bear children in order to fulfill a “moral duty,” which adds to the stigma surrounding abortion in Ireland.\textsuperscript{39}

Yet another aspect of daily life in Ireland that contributes to the country’s restrictive policy is the persistent silence of women who have had abortions, as Ruth Fletcher examines in her paper \textit{Silences: Irish Women and Abortion}. In the paper, Fletcher shows that, while abortion is a complex issue, the general public is presented with a polarized debate that demonstrates two distinct and opposing sides. As Fletcher states:

\textit{The pro-choice movement has been associated with a perception of abortion as a straight-forward procedure which a woman undertakes in pursuit of control of her reproductive capacity with little or no consideration of the fetus. The image of abortion that is connected with the pro-life movement is one of an evil act where that woman is responsible for the killing of an innocent unborn child, resulting in her feeling guilt and remorse.}\textsuperscript{40}

However, this debate does not include the opinions of women who have actually had abortions, none of whom, according to Fletcher, “identified with either of the two oppositional depictions of abortion”
presented to the public. Many of the women interviewed by Fletcher state that the main reason for their silence was that they wanted to avoid the harsh criticism and disappointment directed at them from their society. Each of them felt that if their secret was known, they would be “viewed as irresponsible people…no matter how responsible they felt in making their decision.”

Japan

The abortion policy in Japan exemplifies the limited level of restrictiveness. According the Penal Code of Japan, abortion is illegal. However, doctors are able to approve abortions under certain conditions laid out in the Maternal Health Protection Law of 1996. These conditions include circumstances in which the continuation of the pregnancy would endanger the life and health of the mother and if the pregnancy was the result of rape.

Despite the illegal status of abortion, a survey conducted in 1998 showed that “79% of unmarried and 85% of married women approved of abortion…. [and that] nearly one-third of all abortions are performed on women younger than 25.” This trend can be explained by the fact that while the acceptance of premarital sex has increased steadily over time, the “social acceptability of childbearing outside of marriage” has decreased, resulting in an increased number of abortions for young, unmarried women.

At first glance, the restrictive policy of the Japanese government seems to contradict the attitudes of the general public. However, this apparent inconsistency can be explained when looked at in a religious context. On the whole, Japan is a Buddhist country and abortion, according to Buddhism, is an act that causes suffering. This concept is thus reflected in the official policy. However, as previously stated in this paper, the idea of karma also plays a key role in the decision making process of Buddhists. Therefore, the general public is more open to the idea of abortion because they simply weigh the karmic cost of abortion with other socio-economic factors.
Thailand

Abortion regulation in Thailand is very similar to that in Japan. Article 305 of Thai Penal Code makes abortion illegal except in cases when it is deemed appropriate by a doctor. The grounds upon which abortion is most often permitted are in order to save the woman’s life, preserve her physical and/or mental health, and when the pregnancy is the result of certain sexual offenses, such as rape and incest.44

Also similar to Japan, there is a range of socio-economic reasons that women list as their reasons for undergoing illegal abortion procedures. These factors include understanding of fetal development, gender relations, class position, and notion of sin. According to Andrea Whittaker, such factors provide “justifications for abortion that are congruent with broader social goals and expectations.” Just as in Japan, women in Thailand are generally willing to consider abortion, despite its illegal political status, favoring the resulting negative karma over a much more difficult life in the long term. As one woman Whittaker interviewed stated: “if you talk about the issue of sin, [abortion] is a sin, right. But we first have to think about the truth of our day by day lives, whether we can accept it or not according to how we live.”45

Iran

In recent history, policies concerning abortion regulation in Iran have altered with each change in political regime. Abortion was first legalized in 1978 under the rule of Mohammad Reza Shah Pahlavi, a leader installed by the United States’ CIA who was very sympathetic to Western ideas. When Reza Shah Pahlavi was overthrown in 1979, Iran came under the rule of Ayatollah Khomeini, a religious leader who set up a theocratic Muslim state disapproving of abortion.

More recently in 2005, a new bill was approved that altered the conditions required to obtain a legal abortion. This categorizes Iran’s abortion regulations as limited. These conditions include if the mother’s life is endangered and if the fetus shows signs of handicap. Additionally,
the mother must receive consent from three physicians and the abortion must take place with the first nineteen weeks of pregnancy.\textsuperscript{46}

Iranian policy on abortion reflects the value Islam places on the life of the mother. While the fetus is indeed considered a person after ensoulment, until it is born, the life of the mother always takes precedence over the life of the child. Thus, while there are certain regulations put in place to ensure that the abortion is necessary, the procedure is always permitted in order to preserve the health of the mother.

\textbf{Israel}

In Israel, abortion is considered legal under certain circumstances. As such, Israel’s abortion policy falls under the category of \textit{conditional}. In order to get an abortion in Israel, a committee for pregnancy termination, or Termination Committee, must approve of the abortion under the following conditions: the woman is younger than 17 (the legal marrying age) or older than 40, the pregnancy is the result of illegal activity, such as rape, incest, or adultery, the fetus shows signs of handicap, either physical or mental, or the continuation of the pregnancy would turn out to be harmful to the mother’s health.\textsuperscript{47}

There are numerous conditions under which a woman in Israel may obtain an abortion. This relates to the principles of Judaism because the fetus is not considered an individual until it is born. Therefore, the life and well-being of the mother is always considered of more value than the preservation of the fetus.

\textbf{Finland}

Like Israel, Finland also has a \textit{conditional} policy on abortion. In 1950, the Parliament of Finland first legalized abortions, allowing them in cases where the mother’s life or health was in danger, if the woman was under 16 years old, if the fetus was deformed, and if the woman had been raped. Twenty years later, in 1970, the Parliament revised their previous ruling, setting a time limit on when legal abortions could be
received. Under this ruling, women who were younger than 17, older than 40, or who already had four children were eligible for abortion up to 16 weeks after conception. Nine years later, they revised the law again, lowering the 16 weeks to 12 weeks. In 1985, Parliament changed the time limited once more, allowing underage women 20 weeks to get an abortion, 24 if the ultrasound or amniocentesis showed potential defect in the development of the fetus. Additionally, a woman must get the approval of two physicians before she can receive an abortion. The only exception to this rule is if the woman is underage, overage, or already has four children—in these cases, she only needs the approval of one physician.48

The conditional abortion policy in Finland relates to the country’s religious atmosphere. Unlike the populations of the other countries examined in this paper, the people of Finland do not practice a religion found worldwide. Rather, the most prominent religion found in Finland is the Evangelical Lutheran Church of Finland. While certainly grounded in Christianity, the Church of Finland is unique to that country and its constituents. As such, there is more flexibility for the government to create policies that are agreeable with the Church and, by extension, the particular people the policy will ultimately affect.

Federal System

The final two countries that this paper will examine are Australia and the United States. However, it is impossible to categorize them using the same very restrictive, limited, and conditional because they operate under a federal system. In a federal system, the country as a whole is split into a number of states or provinces, making it so that each constituent is ruled by two distinct governments: the individual state government and the national (or federal) government. In the case of both Australia and the United States, abortion is legal under federal law, but the state governments have leeway to restrict and regulate the places in which abortions may be offered and the people who are licensed to perform abortions. These state regulations can severely limit abortions, even if they technically remain legal. Thus, it is impossible to categorize the country as a whole.
Throughout Australia, constant in every state, abortion is legal in cases where it is deemed necessary to save the life and/or health of the mother. However, this definition changes state to state. For example, in Queensland, to preserve the physical or mental health of the mother is the only situation in which a woman may receive an abortion. In the Northern Territory, South Australia, and Tasmania, a woman can also get an abortion if she is the victim of rape or the fetus shows early signs of defects. In the territories of New South Wales and Western Australia, the policy becomes more liberal still, since the woman is allowed to apply for an abortion based on both social and economic factors as well.49

In their article Abortion and Health Care Chaplaincy in Australia, authors Lindsay Carey and Christopher Newell discuss how abortion, like in many other countries, has been a “long standing polarized bioethical issue within Australia.”50 Carey and Newell lay out the three central Australian perspectives on abortion: the conservative, moderate, and liberal positions. Depending on how the population is dispersed, each of these positions dominate the policy in at least one Australian territory. Similarly, due to the lack of religious unity in a single territory, religion does not play as large a role in policymaking in Australia as it does in countries with a single, central government.

By its very definition, the federal system does not provide much unity when it comes to policies on issues such as abortion. As Carey and Newell discuss, one solution that has been proposed to solve the debate between the three ideological positions surrounding abortion is the one abortion policy, that is “each woman should be legally permitted to have one abortion during her life. Thereafter the rights of the fetus prevail… except where the pregnancy arose out of forced sex or its continuation presents a demonstrable and serious risk to the welfare of the woman.” Not only would this solution appeal to a wider audience than a more polarized policy, the one abortion policy would also save numerous lives by reducing the number of abortions while at the same time still allowing women a certain measure of control over their own lives.51
United States of America

The current state of abortion policy in the United States is contradictory. On the one hand, the overarching policy of the federal government is based on the decision made in the court case *Roe v. Wade* (1973), which established women’s right to abortion. On the other hand, a significant number of state policies tend to follow the decision made by a later Supreme Court case *Webster v. Reproductive Health Services* (1989), which enacted regulations on how individual states may assist with abortion procedures.\(^\text{52}\) Thus, in short, the United States operates under the judicial interpretation that abortion is legal, yet may be regulated by state governments.

In the years since the 1973 *Roe v. Wade* case, the individual state governments have exercised their freedom and implemented various policies concerning regulation on abortions. These regulations tend to differ based on the dominant ideology of the state’s constituents. For example, in California, a state dominated by secular, liberal ethics in Los Angeles and San Francisco, the law does not require women seeking abortion to undergo any of the usual restrictions, such as waiting periods, limitations on funding, or parental involvement (in the case of pregnant minors).\(^\text{53}\) In other states, such as Texas, which is dominated by Conservative Christian constituencies, abortion policy is much more restricted. In Texas, there are only six legal abortion clinics in the entire state, which means that some women have to travel hundreds of miles in the first twenty weeks of their pregnancy to have access to a clinic.

As stated above, the variety of state abortion policies is based on the differences in ideology, and it is important to note that ideology can be heavily influenced by religious beliefs. Indeed, the American South, or “Bible Belt” as it is sometime referred to, is widely regarded as one of the most religiously conservative parts of the United States. Religion has been a long-standing player in the debate over abortion in the United States. In 1976, Congress held a hearing on a proposed amendment to the Constitution that was aimed at overturning the *Roe v. Wade* decision. During this hearing, members of Congress heard
representatives from both the Roman Catholic Church, a group that is
determined to keep abortion illegal, and the United Methodist Church,
which had recently joined a coalition to defend the *Roe v. Wade* ruling.
While the amendment did not pass, the representatives from both
parties guaranteed they would “mobilize their constituents on behalf
of their perspectives.” Since then, the various denominations of
Christians, as well as a few representatives from other religions, have
made sure their voices are heard in the ongoing debate over abortion.

While such religious groups have not made a difference on the federal
abortion law, they have influenced state laws across the country.
It is these religious groups that are primarily responsible for the
contradictory nature of American abortion policies. As Yael Yishai
states in his paper *Public Ideas and Public Policy: Abortion Politics in
Four Democracies*, “[while] a state acknowledges the right of a private
person to make her own reproductive decisions…[it] is not committed
to enabling her to act upon this decision.” This conflict between
individual choice and lack of state resources has proved to be harmful
to many women, especially those who depend on public aid.

Conclusion

Most societies, even authoritarian ones, legitimize themselves by saying
that they defend the rights of their citizens. With abortion, governments
have had to confront the conflict of competing rights: a woman's right
to bodily integrity vs. the fetus’ status. Most modern governments are
not theocracies—that is, governments tasked with enforcing a certain
religious viewpoint—and even in those societies, religious viewpoints
may differ. On the whole, based on the evidence outlined in this paper,
all governments seem to strive to balance religious sensibilities with
modern scientific information: permitting abortions early in gestation,
and imposing more limits and restrictions as pregnancy advances.
Interestingly, even some restrictive governments permit abortions in
the case of fetal abnormality: a throw-back to the Roman idea that less-
than-normal infants were not entitled to the status and protection of
citizens. This is becoming another sensitive issue in the United States,
where Americans with disabilities have definite protections under
federal law. Anti-abortion activists increasingly resist abortions even in the case of gross abnormalities in which the fetus would be non-viable.

Furthermore, the pragmatic arguments raised in Asian countries—that while abortion is associated with negative karma, it also is balanced against the long-term effects of continuing an unwanted pregnancy (un-loved or neglected or rejected child, or negative socio-economic effects on the mother and mother’s family)—play a role in women’s decision making process worldwide. Such pragmatism, however, is not voiced as clearly in Western countries because it doesn’t rest as fluidly in political discourse as does the issue of legal rights of individuals. In the United States in particular, abortion is not seen publicly in this context, no matter how much individuals, in the confines of their family circles, may try to articulate and accept the complexity of abortion.

Notes


3. Ibid.


5. Ibid.


10. Ibid.

11. Ibid.


17. Ibid.

18. Ibid.


20. Ibid.


24. Ibid.


26. Ibid.


29. Ibid.


34. Brahm, “When Does Human Life Begin?”

35. Whittaker, “The Struggle for Abortion.”


38. Ibid.


40. Ibid.

41. Ibid.


50. Lindsay B. Carey and Christopher Newell, “Abortion and Health Care Chaplaincy in Australia,” *Journal and Religion and Health* 46, no. 2 (June 2007).

51. Ibid.


55. Yishai, “Public Ideas and Public.”
A Faulty Cause and Effect: Exploring the Fallacy of the Connection Between Dred Scott and the 14th Amendment

Katherine Keller

On March 6, 1857, Supreme Court Chief Justice Roger B. Taney issued the decision of the seven to two majority in the case of Dred Scott v. Sandford. Taney’s racist rhetoric, simultaneously declaring that congressional regulation of slavery was unconstitutional and that free Blacks were never intended to be and never could become citizens, created great controversy in a nation already plagued by sectional tension. Historian James Kettner argues that Dred Scott had a transformative impact on the movement for Black citizenship, bringing the issue to center stage and widening the ideological disparity between North and South as the Civil War approached. However, Don Fehrenbacher and other scholars hold that the movement for Black citizenship evolved separately from the reaction to the Dred Scott decision, as the groups who advocated for Black citizenship formed too small a constituency to pass even the 13th Amendment by themselves, let alone the more controversial 14th. Radicals and intellectuals, those who formed this constituency, needed the support of the moderate Republicans in order to make Black citizenship a reality. However, while moderates opposed the parts of Taney’s decision that struck at the Free Soil ideology, most of them agreed, explicitly or implicitly, that Blacks should not be citizens.

It took even more radical changes than Dred Scott to unite Republicans behind the movement for Black citizenship; such changes sprang forth from the revolutionary Civil War. The changing aim of the war, as characterized by the Emancipation Proclamation of 1863, ideologically united the North behind the doctrine of Blacks’ human rights. From there, Blacks’ contribution to the war effort, coupled with the hostile environment Blacks faced in the South during the early years of Reconstruction,

This paper was written for Dr. Charles Hanson’s Advanced Placement U.S. History class in the spring of 2014.
finally convinced the majority of moderate Republicans to support the civil rights for which Radicals had advocated even before 1857.4

Backlash Leads to Change?

It is true that Taney’s ruling against Black citizenship created a considerable amount of backlash among certain groups and motivated such constituencies to advocate for a change in policy. The most prominent group in opposition to the citizenship clause of Dred Scott was the Radical Republicans, who composed a loud minority of the North. In the week following the decision, numerous abolitionist and Radical newspapers throughout the North expressed their outrage at the explicit disenfranchisement of Black people that Taney articulated in the decision of the court. Such editorials served as a call to action to overturn the Supreme Court’s “barbarous and humiliating” decision against Black citizenship.5 Radicals’ enrage against Dred Scott was not limited to proximate reactions to the decision, either; prominent abolitionist William Lloyd Garrison published an editorial in his abolitionist newspaper The Liberator a year later that contained rhetoric just as impassioned as that in editorials published the day after Taney’s decision.6 Radicals’ unwavering advocacy for Black citizenship in protest of the Dred Scott decision formed one of the first movements in favor of civil rights for freedmen.

The Radical Republicans’ protest was joined by a large part of the intellectual community, who found flaws in the logic that Taney used in his ruling against Black citizenship. Benjamin Howard’s scholarly article published in The North American Review in the months following Taney’s decision exemplified intellectuals’ criticism of the poorly wrought logic in the citizenship clause of Dred Scott. Using the same strict-constructionist viewpoint as Taney did in his decision, Howard discredited Taney’s claim that the framers had never intended for Blacks to be citizens, using examples of five states that had granted civil and political rights to Blacks when the Constitution was ratified.7 Intellectuals like Howard differed from abolitionists like Garrison in the rhetoric they used to criticize Dred Scott, but from an ideological standpoint, the two groups were united by their shared belief that Taney’s ruling
against Black citizenship needed to be overturned. For these two constituencies, the Dred Scott decision had raised the cause of Black citizenship to a central focus and motivated them to make the first steps toward what would become the 14th Amendment.

Section 1: Moderates Oppose Black Citizenship

While the Dred Scott decision did unite radical Republicans and intellectuals from an ideological standpoint in favor of Black citizenship, their voices alone were not enough to bring their cause to fruition; they needed the support of moderate Republicans to pass Black citizenship. The two major pieces of Reconstruction legislation pertaining to Black citizenship, the Civil Rights Act of 1866 and the 14th Amendment (first passed by Congress in June of 1866), were passed during the Thirty-Ninth Congress, in which Radicals had to cooperate with a moderate majority. It was not until the elections in November of 1866 for the Fortieth Congress that Radicals gained complete control over Congress.  

Moderates generally opposed Radicals’ endorsement of Black citizenship before the war; in fact, the Dred Scott decision solidified most moderates’ opposition to Black citizenship. Most Republicans, radical and moderate, agreed that slavery should not be allowed to expand, but abolitionists’ and moderates’ views on the civil rights of freedmen sharply differed. To the dismay of abolitionists, moderate Republicans sought to revive their appearance as “the white man’s party”, an image that became more and more threatened as the abolitionist movement strengthened, by voicing their opposition to Black citizenship. During his 1858 campaign, Lincoln exemplified and led the movement in favor of Free Soil but fervently against Black citizenship; many of his fellow moderates aligned their beliefs with the ones for which he advocated during the Lincoln-Douglas debates. Thus, the backlash against the Dred Scott decision that advocated for Black citizenship was limited to abolitionists and intellectuals, and those two groups did not make up a large enough population to actually accomplish Black citizenship without the help of the moderates.
Moderate Republicans did oppose the Dred Scott decision vehemently, though their opposition to the case mostly concerned the part of Taney’s opinion that directly struck at the Free Soil ideology. Apart from his ruling against Black citizenship, Taney’s opinion also included the declaration that the Missouri Compromise was unconstitutional. Taney argued that Congress did not have the power to pass a law that would prohibit slavery in a specific territory, which was exactly what Free Soilers sought to do in the West. Thus, Taney had essentially declared the main platform of one of the two major political parties to be unconstitutional. Unsurprisingly, many Free-Soil moderate Republicans took great issue with this ruling. Directly following the case, Republican newspapers such as *The New York Daily Times* harshly criticized this aspect of the Dred Scott decision, lamenting the recent “nationalization of slavery.” However, such opinions avoided the more controversial topic of Black citizenship.

Another popular argument among moderate Republicans against *Dred Scott* was the idea that Taney’s decision was *obiter dictum*, meaning that once the Supreme Court decided that it did not have jurisdiction in the case, everything it said thereafter was not authoritative. This argument extended the Republican opposition to the anti-Free Soil part of the decision by articulating a way around it; while they believed that the decision should one day be overturned, critics who favored the *obiter dictum* argument held that in the meantime, it could be dismissed as not authoritative. One editorial from *The New York Tribune* exemplified this belief, stating that the ruling against the Free Soil ideology held no more weight than “the judgment of a majority of those congregated in any Washington bar-room.” Such opinions at first appear to align with Radicals’ beliefs, but the *obiter dictum* argument against the Dred Scott decision implicitly opposed Black citizenship. The Supreme Court rejected the case because it did not have jurisdiction, since Dred Scott was not a citizen of the United States and therefore could not sue. The *obiter dictum* argument rested on the assumption that the Supreme Court indeed did not have jurisdiction in the case, and by extension, that Dred Scott was indeed not a citizen.
Section 2: Changing War Aims Lead to Changing Northern Ideology

Only the Civil War provided the revolutionary landscape necessary to shift moderates’ ideology from Free Soil to one in favor of emancipation and civil rights for Blacks. During the first year of conflict, however, many moderate Republicans and northern Democrats still clung fervently to the anti-Black sentiments that had separated them from the Radicals during the antebellum years. The North could unite, however, behind the goal of preserving the Union, which remained the official aim of the war during its first two years. Lincoln expressed this singular goal in the famous letter to Horace Greeley in which he stated that despite his “personal wish” for emancipation, his focus in the war was to preserve the Union at all costs, regardless of the fate of slavery. Many who fought in the Union army echoed this focus on preserving the Union, the “beacon light of liberty and freedom to the human race.” Even Lincoln’s preliminary Emancipation Proclamation of 1862 gave the South the option of surrender in exchange for the continued institution of slavery.

As the war continued, however, even pragmatic moderates began to support emancipation. The ideological shift in the aim of the war, as characterized by Lincoln’s Emancipation Proclamation, served as an important step toward achieving civil rights for Blacks because it replaced the Free Soil ideology with unhindered abolition. Uniting the North behind the goal of emancipation, as opposed to just the preservation of the Union, rallied support for at least the human rights of Blacks. Without unanimous support for Blacks’ human rights, advocates of Black citizenship could never have achieved civil rights for freedmen. As Eric Foner illustrates, a population can be enfranchised only by working their way sequentially up the ‘pyramid’ of rights: gaining first human rights, then civil rights, then political rights, then social and economic rights. As slaves, Blacks had been stripped of human rights—the lowest tier; before their civil rights could be considered, a majority of the North had to support abolition, the movement for their human rights.
The change in war aim from union preservation to emancipation was widely viewed as a necessity, from both humanitarian and military standpoints. Lincoln's Secretary of the Navy, Gideon Welles, wrote that just before issuing the Emancipation Proclamation, Lincoln expressed his conclusion that emancipation was “absolutely essential for the salvation of the Union,” and would benefit the Union military effort despite the seemingly inevitable backlash from slave owning border states that had remained loyal to the Union. Even moderates who had previously opposed the abolitionist movement could agree that emancipation was, strategically, a good military tactic. As Henry W. Halleck wrote to Ulysses S. Grant in March of 1863, emancipation would transfer labor to the Union by enabling the use of slaves who came under Union control as soldiers. Emancipation provided many more Blacks with the opportunity to fight for the Union Army; Blacks made up ten to twelve percent of the Union’s military manpower by the end of the war.

The experience of fighting in the South also shifted sentiment toward abolition; for many Union soldiers, the war was the first time they had ever witnessed the humanitarian horrors and the backward nature of slavery firsthand. For example, Private James Miller of the 111th Pennsylvania Infantry displayed the newfound disgust with slavery that many of his contemporaries shared; while he had never demonstrated abolitionist sentiments before the war, he wrote to his brother that the abysmal treatment of slaves he had observed in Virginia had solidified his utmost support for emancipation. Furthermore, after witnessing an elderly slave woman risking her life to run away with her granddaughter after the baby’s mother had been sold, Private Constant Hanks of the Twentieth New York Militia wrote that the dissolution of slave families that occurred throughout the South “would wring the tears out of anyone’s eyes.” Thus, many soldiers had already come to oppose slavery even before the Emancipation Proclamation.

The redefinition of the Union’s war aim was not received with unanimous support at first, but by the end of the war, the vast majority of the North agreed that Blacks deserved human rights. Private Chauncey Welton of Ohio exemplified the ideological shift that occurred in even the most conservative supporters of the Union. At first, the Northern
Democrat lamented that the Emancipation Proclamation had changed the aim of the war to one for which he was not willing to fight. However, he began to view emancipation as a necessary step toward stopping the rebellion, and even had enthusiastically joined the Republican Party by the end of the war. By 1864, Lincoln’s reelection campaign in favor of a constitutional amendment abolishing slavery earned him almost eighty percent of the soldier vote.

This great shift in public opinion in the North in favor of human rights for Blacks constituted an important step toward Black citizenship. In order to achieve civil rights, Blacks had to achieve human rights first, and the Civil War fostered the ideological revolution in the North that allowed these human rights to be legally recognized.

Section 3: Fulfilling the Duties of Citizenship and Upholding Emancipation

With the doctrine of emancipation, and thus human rights for Blacks, secured ideologically among the vast majority of the North, the movement for civil rights and enfranchisement of freedmen could begin to take root among moderate Republicans and northern Democrats. Just as the doctrine of human rights did not necessarily accompany the antislavery ideology (as the antebellum Free Soil movement exemplified), the doctrine of civil rights did not necessarily accompany emancipation but rather emerged because of other factors. However, the doctrine of civil rights could not have gained traction without the base that abolitionist ideology provided. One factor that influenced moderates’ support for citizenship and civil rights of freedmen was Black service in the Union Army during the Civil War. Following the Emancipation Proclamation, Black activists including Frederick Douglass foresaw that Black service in the military would help secure their civil rights following the war because it would show that they could fulfill the duties of citizenship. This ‘call-to-arms’ to northern freedmen and newly emancipated slaves in Union-controlled territories proved to be duly beneficial; Black regiments contributed greatly to the Union victory, and Black service did end up playing a key role in earning the respect necessary to secure civil rights for their race.
During congressional debate of the Civil Rights Act of 1866 (the revolutionary legislative precursor to the 14th Amendment), former Union generals such as Benjamin Butler made impassioned speeches that advocated for Black citizenship through anecdotes of their heroism on the battlefield. Moderate Republican congressman and former general John Farnsworth similarly declared that “good faith, as well as impartial justice, demands of [the] Government that it secure to the colored soldiers of the Union their equal rights and privileges as citizens of the United States.” Such sentiments of respect and gratitude for the sacrifices that freedmen made on behalf of the Union, and the citizenship that they deserved as a result, reverberated throughout the moderate and radical factions alike and even were implicitly echoed by President Lincoln himself in his last public address, when he voiced his wish that educated freedmen, and those who had served as soldiers, could enjoy the right to vote (a right that would require citizenship to exercise). To much of the North, Blacks’ service in the Union Army proved that they were willing and capable of fulfilling the duties of citizenship.

Another factor in moderates’ ideological shift to supporting Black citizenship was the South’s attempt to restore the legal, social, and economic hierarchy that had existed during the antebellum period. Unwilling to make such rapid and revolutionary changes to their longstanding society, southern states passed Black Codes in a subversive effort to continue restricting Blacks’ legal rights and economic opportunities. Mississippi and South Carolina initiated such laws that violated free labor precepts in late 1865, and the rest of the South followed suit. Among other restrictions, these laws prohibited Blacks from renting land, required them to sign labor contracts that surrendered their right to strike, and prohibited them from leaving jobs once contracted. Moderate and radical Republicans alike wanted to establish a free labor economy in the South, and the institution of the Black Codes convinced many moderates that more legislative action was needed in order to do so. For example, during congressional debates concerning the Civil Rights Act of 1866, which would grant citizenship and civil rights to freedmen, moderate Republican leader Jesse Fell of Illinois advocated for the necessity of “[adopting] measures for the safety and elevation of the African race” in order to mend the “mockery” of “their present nominal freedom.”
Furthermore, Lincoln’s close friend Gustave Koerner objected to southern states’ omission of freedmen’s civil rights in their newly-drafted constitutions, lamenting the “most obnoxious laws . . . [which] placed most of the free negroes under a sort of Mexican peonage.”

Thus, moderates’ respect for Black soldiers’ sacrifice in the war coupled with their sense of duty to reverse Black Codes in the South enabled their ideological union with the Radicals on the issue of Black citizenship in the few years after the war’s end. Now, with most of the North supporting freedmen’s civil rights, and the citizenship they needed in order to enjoy them, it became possible to pass legislation on the issue. Despite President Johnson’s veto of the preliminary Civil Rights Act of 1866, the Radicals and the moderates together proved to be a large enough constituency to both overturn the veto and pass the 14th Amendment in 1866 (which would be ratified two years later). The first section of the amendment declared that everyone born or naturalized in the United States was a citizen, regardless of race. Thus, the citizenship clause of the Dred Scott decision finally had been overturned, but as a result of many factors other than backlash to the case itself.

Looking Ahead: Enforcement and Precedent

It is true that the 14th Amendment was difficult to enforce in the South, where institutions such as sharecropping, the Ku Klux Klan, and Jim Crow laws upheld the ingrained social hierarchy of antebellum times and restricted the civil rights that the Reconstruction amendments conferred on Blacks. However, the adoption of the 14th Amendment set a revolutionary legal precedent for Black citizenship that formed the base upon which freedmen and their descendants began to enfranchise their race. The racial environment in the United States could not be reversed de facto with the adoption of three amendments, but without them the long journey to Black empowerment never could have begun.

_Dred Scott_ is one of only a handful of Supreme Court cases to have been completely overturned by constitutional amendment, and many modern politicians wonder if the controversial decision in the case of _Roe v. Wade_ (1973) might follow suit. Like the Radical Republicans in 1857,
conservatives today are enraged that the Supreme Court ruled against the human rights of unviable fetuses (less than twenty-eight weeks), calling *Roe v. Wade* “virtually identical” to the “dreadful opinion” that Taney published. Conservatives’ impassioned objection to *Roe v. Wade* and their advocacy of a constitutional amendment overturning it have made abortion a highly debated issue for the past three decades. However, the precedent of *Dred Scott* and its reversal by the 13th and 14th Amendments, if the connection between the two articulated in this paper holds true, makes the adoption of an anti-abortion constitutional amendment seem unlikely. Similarities between the two cases definitely exist, but the great disparity between the two situations consists in the political climates in which each took place. To rally a large enough constituency to reverse *Dred Scott*, it took the truly revolutionary environment of the Civil War; despite the controversy over *Roe v. Wade*, the revolutionary environment required to overturn the case does not yet exist. If the historical precedent of *Dred Scott* and the Reconstruction Amendments holds true in modern society, conservatives advocating for the adoption of a pro-life amendment will most likely remain a loud and aggrieved minority.

Notes


23. Pvt. James Miller to his brother, August 17, 1862, in *What This Cruel War was Over* by Chandra Manning (New York: Knopf, 2007), 77.

24. Pvt. Constant Hanks to his mother, August 8, 1862, in Manning, *What This Cruel War was Over*, 76.


26. Ibid., 65.

27. Ibid., 67.


34. Ibid.


Bibliography

Primary Sources


The Liberator (Boston, MA), March, 1858.


Secondary Sources


Where Practicality Meets Idealism: Reasons for the Criticism of the Alaska Purchase

Joseph Klein

On March 30th, 1867, Secretary of State William Seward and Russian foreign minister Eduoard de Stoeckl agreed to a treaty that ceded the Russian America territory to the US in exchange for $7.2 million. After months of bickering over the initial terms of the treaty in both the Senate and the House, Alaska, as it was called, was formally transferred to the United States in December of 1867. Russia had previously approached the United States about selling the territory in 1859, but the Civil War delayed the negotiations between the two nations. When the treaty was inked in March of 1867, very few members of the American public had any idea that the United States was attempting to acquire Alaska, which implies that very little factual information regarding the territory was available, especially to newspapers located on the East Coast. The acquisition of the largely unknown piece of land left the American public scrambling for facts and opinions because the purchase had come about so suddenly. The debating and bickering continued into the summer of 1868, when negations in the House of Representatives regarding the appropriation to Russia for Alaska were finished and the deal was finalized.

A change of this kind can often lead to backlash, whether or not the challengers of the new policy or program are in the majority. In the case of the Alaska purchase, the majority of American newspapers were in favor because of the assertions of its economic potential as, well as its strategic location for military operations and possible Pacific dominance. But epithets such as “Seward’s Icebox,” “Johnson’s Polar Bear Garden,” and “Walrussia” did not come from the purchase’s supporters. The critics, despite being a minority, were a very loud group, and they publicized their opposition to the concept of Alaska becoming a United States territory.

This paper was written for Dr. Charles Hanson’s Advanced Placement U.S. History class in the spring of 2014.
The criticism was loud, but why? What got some newspaper editors and congressmen so irritated about the Alaska purchase? Historian Thomas A. Bailey argues that the negative sentiment surrounding the Johnson administration led to the negative reviews of the treaty. The criticism was twofold: some were upset about Johnson and his Reconstruction policies, while others lamented Seward's perceived upward management and his greed regarding territorial acquisitions.

But that answer does not represent the majority of the critics. While many were seemingly discontent with the White House's actions, there was something about Alaska that got the critics really upset. Alaska, in most critics’ minds, was just a bad business deal, both in the short and in the long term. Proximately, Alaska seemed to have very little value, and what value it did have would never be harvested because nobody would want to move there due to its notoriously bad weather. Ultimately, pragmatists pointed to the money it would take to defend a noncontiguous territory and maintain (as well as set up) a functioning government capable of keeping law and order and peace in the territory as the natives clashed with the US military in 1868 and 1869. While Alaska now seems like the best business deal in American history, it was seen by many as a high risk, low reward purchase that was not worth it.

**Dissent Directed at the Johnson Administration**

The Johnson Administration had done itself no favors with much of the public—notably the Radical Republicans—when it tried to get the Alaska treaty purchase through Congress. Newspaper editors used the Alaska Purchase as a channel to criticize the current government, a critique that took on the President and his Secretary of State.

President Andrew Johnson's policies towards Reconstruction and the new South were not well liked; he was impeached just a year after the Alaska Purchase treaty was proposed. However, at the time of the Alaska treaty, “the Reconstruction controversy...had just entered its most acute phase.” President Johnson was starting to feel the heat from both the Radical Republicans, who wanted tighter, tougher restrictions on the
white South, and the white South, who wanted liberal laws towards the “new” Southern states. At a time when resentment was growing, Bailey argues that “the Alaska treaty was regarded as not only an annoying interruption but as an effort on the part of the harassed administration to cover its blunders at home with a spectacular triumph in the field of diplomacy.”

Others looked through the same lens and saw Johnson purchasing Alaska as a cover up for the elections that were coming in 1868. At the time, nobody could have predicted that Johnson would be impeached, thus ending his chance at getting re-elected. But Frank Leslie, in his cartoon *Preparing for the Heated Term* (see Appendix D), believed that Johnson was using the purchase, which was actually quite popular, as a method to regain political legitimacy by passing a favorable bill. Leslie was a staunch Radical Republican who attacked Johnson and Seward whenever he had the chance (see Appendix B). Bailey agrees, believing that many Radical Republicans did not want to see Seward and Johnson get credit for anything popular.

Johnson’s Secretary of State, William Seward, was also under fire for the Alaska purchase. Many saw him as an upward manager who would do anything to get on Johnson's good side (see Appendix A). Likewise, the *Evening Telegraph* from Philadelphia asserted that “[Seward] saw money in [the Alaska Purchase], not merely for the indefinite American citizen, but for William H. Seward in particular.” This critique, which came after Seward’s term as Secretary of State came to a close, suggests that Seward had private financial goals in mind when purchasing Alaska, further angering some Americans who were enraged that their tax money was going into the pockets of an already wealthy man. Seward was not only seen as greedy in wanting political and economic power, but also for acquiring territory. The *Charleston Daily News* believed that “territorial expansion [was] Mr. Seward’s hobby.” Staff writers of the paper also thought that Seward, despite it being his job description as Secretary of State to execute America’s goals abroad, was too aggressive in his “plots…to extend the American Union to the Arctic Sea.” Despite rampant Manifest Destiny, which had existed in America since colonists refused to accept British
demands that colonists stay east of the Appalachians, many Americans disapproved of Seward’s lust for territorial expansion. According to the *Evening Star* in Washington DC, Seward was seen to have purchased “icebergs, tornadoes, and earthquakes,” the same way the Louisiana Purchase was assailed in 1803. Some Americans were tired of trying to expand the country when there were serious problems at home that needed to be dealt with. Seward was the wrong man for the job during this time: as an idealist who wanted to create an American empire, he was labeled by pragmatists as someone who did not care about American affairs and just wanted to pursue his own interests, kissing up to Johnson along the way to get the funds necessary.

Anti-Johnson and anti-Expansionist movements, however, do not provide a satisfactory explanation for criticism. When it came time to vote on the appropriation bill in 1868 to pay for the purchase, some of Johnson’s staunchest opponents were willing to cross party lines and vote in favor of making an appropriation. Thaddeus Stevens, a staunch opponent of Johnson and an integral member of the impeachment movement, was supportive of the administration’s plan for paying for the bill. This pro-Alaska stance from a radical Republican who loathes Johnson and everything Johnson stood for could be attributed to the release of new information by July of 1868. Whereas the territory was largely unknown at the time of the purchase in March 1867, information flooded into the contiguous states in the months following, allowing more newspaper editors and politicians to make informed opinions. There was something about Alaska, at the time of purchase, which really ruffled the feathers at many newspapers.

**Initial Concerns: The Barren Wasteland**

Immediately after the purchase was made, uninformed newspapers criticized the initial $7.2 million sum that the United States had to pay to Russia for Alaska, believing that Alaska had very little economic value and that the minimal value it had would never come to fruition because no person in their right mind would move there. Many Americans believed that Alaska had no economic value because the Russians were so willing to sell the territory to the USA. At the
time, oil had not been discovered in the territory; the most valuable trade item was believed to be fur. As the *New York World* put it, “that [fur] trade has declined and nearly run out by the destruction of the animals.” Russia had sold the US a territory whose economic value had been squeezed out, from these critics’ point of view. In hindsight, this incongruity can be attributed to a lack of good information about the territory, as very few reporters were willing to go to Alaska to find out more. Most of the criticism came from East Coast newspapers that had no interest in learning more about the potential economic value of the Russian territory. Thus, it was accepted amongst the critics that “Russia [had] sold us a sucked orange.”

The topography of Alaska was largely unknown to the newspaper editors throughout the country, making it easy for them to assume that Alaska was filled with “impassable deserts of snow… [and] inaccessible mountain ranges.” As Missouri’s *Holt County Sentinel* assumed, all the US got by acquiring the territory, was the city of Sitka, the Prince of Wales Island, and “waste territory.” It was unacceptable to pay $7.2 million, in their minds, for one city and one island that would not have any financial gain when that money could be spent more wisely in the domestic realm. The *Daily Phoenix* in South Carolina sarcastically claimed that “various nostrum vendors had used up all the rocks in this country,” which suggests that these South Carolinians believed that Alaska was all rock and ice and had nothing else. Newspapers, like the *Holt County Sentinel* and the *Daily Phoenix*, that were far away from Alaska tended to share these initial concerns about the geography of the territory, because it was believed that nothing economically beneficial could come out of a land that shared Alaska’s accepted topographical situation. Both of the articles were published in the two months following the purchase, suggesting that accurate information had yet to reach these far away areas.

These initial concerns also included doubts about the weather of the newly acquired territory and how that would affect any economic potential the territory had. Many had caught whiff of rumors about possible mining industries in the territory. However, as Horace Greeley’s *New York Daily Tribune* pointed out, “no energy of the
American people will be sufficient to make mining speculations… profitable.”¹⁹ A good amount of newspapers and congressmen were willing to accept Alaska’s economical potential, with natural resources such as coal, fur, and fish. But what good was economic potential if it could not be turned into economic success? The weather, which was perceived to be subpar at best, was very bad. According to Frank Leslie, a vehement opponent of the Alaska Purchase from the get-go, “there were 107 fair days in the year, and 134 rainy days…. Snow fell on 26 days.”²⁰ The land was tundra, with deserts of snow and no sunshine to be found.²¹ Nobody, these critics assumed, would move to Alaska and help cultivate its resources, and thus it was concerning to many pragmatists who just wanted the United States to benefit from a good business deal.

The dominant attitude on the East Coast which asserted that nobody would move to Alaska was a strong force in politics at the time. Since so many East Coast newspapers described Alaska as tundra with no sunlight, uninformed citizens were not at all excited about the possibility of moving to the former Russian territory. These newspapers believed the purchase was unnecessary and that the $7.2 million could be spent more wisely; they feared that the government’s risky investment would turn into a large liability.

Long-Term Critiques: Security and Setup

For many, however, the issue when it came to the acquisition of Russian America was not the upfront $7.2 million, but rather the long-term financial commitment to a noncontiguous territory. This group saw the Alaska Purchase as an acceptable deal upfront, but believed the deal provided little financial gain and created a hefty long term cost at a time that was by no means a good one for the US government’s balance sheet. As Massachusetts Congressman Benjamin Butler put it in his 1868 speech, “[T]he great question of cost is…what we shall have to pay to take care of it.”²² Butler believed that while the initial cost was simply “a small sum,” the cost to defend the region would be high.²³ There was some immediate truth to the concerns: the US had long been in
contention for areas of the Pacific Northwest with the United Kingdom, even with the Oregon Treaty having been signed twenty years prior. During the early 1860s, Irish Americans, a growing demographic in both numbers and influence in US politics, attempted to invade Canada, further aggravating US-UK relations. Another important consideration was the sheer fact that Alaska was a noncontiguous territory at a time in which the American Navy was not developed to the extent it is today, although it had made significant strides during the Civil War. The only way to get to Alaska by land, at the time, was through British Territory in Canada. That made Alaska vulnerable to an attack during a war with another country, such as with the UK or Russia.24 The amount of money needed to protect Alaska in the immediate being as well as the future made it a military liability, outweighing its potential as a military stronghold for the US in the Pacific.

Another fear of the long-term pragmatist was the cost, in both dollars and time, of setting up and maintaining a government. The cost of creating and operating a territorial government, a judicial system, and police forces was not negligible.25 American politicians themselves were experiencing the difficulty of creating governments when it came to the re-admittance and re-creation of Confederate states. Many critics also realized that there were natives living in Alaska, which creates the issue of tensions with the native Eskimos, as they were called. Butler even admitted that the few American settlers that would go to Alaska would likely “plunder [the natives].”26 Politicians had seen this movie before, with the US having to deal with Native American rights and properties constantly, with battles between the two sides ensuing for the previous fifty years. These problems created stagnation regarding Alaskan development into the 1870’s, leading to more cynicism. An editorial in Frank Leslie’s Weekly Newspaper, a newspaper that was still vehemently against the purchase two years later, criticized the ability of Congress to even begin the government building process, noting “the attempt to erect a territorial government... was squelched in the House.”27 Congress was so focused at the time on Reconstruction policies and impeaching President Johnson that the development of Alaska took a back seat, resulting in neglect.
While Alaska’s civilian government was being built, the US military remained in charge of the territory, leading to a “ruler-subject” relationship similar to the relationship the US had created with the Native Americans during the early 1800’s, leading to much conflict between the military and the “natives.” In the winter of 1869, US garrison buildings in Southeast Alaska were burned down, most likely by native Tlingits. Vincent Colyer, a United States special Indian commissioner, was sent to Alaska in 1869 to inspect the relationship between the Native Americans and the US troops. Colyer, in the annual report of the Indian affairs commission, tellingly said that the “[US troops’] presence tends to demoralize the Indians…. One or the other should be removed.” In the spring of 1870, three years after the purchase, Colyer sent a letter to President Grant describing the violence in Wrangell between US troops and Native Americans. The violence in Alaska that Colyer describes, which was extraordinarily savage on the part of the American troops, made the concerns about Natives and whites living together in Alaska more tangible to the public. The report of the Commission of Indian Affairs, after it surfaced and the public caught whiff of the violence in Alaska, led to more public groaning about the Alaska territory. More money would have to be spent to maintain law and order, because two years after the acquisition, the natives and the soldiers had proven to be incompatible. The prospect of government building in Alaska was a concern at the beginning, but became a big bone of contention as the territory struggled to develop and reports of violence started to leak out.

What’s Next?

Almost 150 years later, it is easy to criticize the critics. Mostly everyone agrees that the acquisition of Alaska for a mere $7.2 million was a superb deal for the United States, especially after oil was found in the territory. Alaska is now third in oil production in the US by states. Gold was found in 1896. The critics have been silenced. Even Seward himself, who had been telling people that generations beyond his would appreciate the purchase, would most likely be stunned at the economic productivity that Alaska has been churning out for decades.
I understand that the exploration of the criticism of the Alaska Purchase cannot prove anything about the nature of backlash in general. But it certainly can add a piece to the puzzle, and that piece is the importance of having accurate information. It is true that Alaska was no “terra incognita,” as historian Richard Welch put it. But it has been shown that the claims about Alaska being a barren wasteland with no inherent value were false. Newspapers did not know, nor did they care enough, about the Russian territory in the Pacific Northwest until taxpayer money was being spent on it. Had more newspapers known about the potential for an Alaskan purchase before March of 1867, we might not have ever chuckled over “Seward’s Ice Box” or “Johnson’s Polar Bear Garden.”

The hardest part about collecting sources for this paper was trying to keep a geographical balance for my newspaper articles. My argument would not be at all convincing if all my sources came from one area, because newspapers in the same area are trying to present facts and opinions to the same audience, often making their arguments very similar. That is why I was so excited to find the Holt County Sentinel, which is from a town in Missouri with a population of 857 as of 2010. I tried to incorporate opinions from places all over the United States, but it was not easy considering the criticism of Alaska was most prominent on the East Coast.

Zachary Jones, a Ph.D candidate at the University of Alaska Fairbanks, is currently studying the violence between Native Americans and the US Army that I described in the previous chapter. Jones’ work is focusing on the Battle of Wrangell in 1869, in which the origin of the fighting is unknown. It was difficult to find anything regarding the violence in Alaska between 1867-1870, so I am looking forward to reading Jones’ work as the study of early Alaskan-American history continues.
Notes


2. Ibid., 44-45.

3. “Preparing for the Heated Term,” in *Frank Leslie’s Weekly Newspaper* (New York), 1867. To see this cartoon, see Appendix D.

4. Ibid.


6. Thomas Nast, “The Big Thing,” in *Harper’s Weekly* (New York), April 20, 1867. To see this cartoon, look at Appendix A. Also of note, more cartoons can be seen in Appendix B.


9. Ibid.

10. *Evening Star* (Washington, D.C.), December 21, 1867. The *Evening Star* was in favor of the Alaska Purchase in this particular article, but the quote that was used was taken out of context. The article was calling for the critics who believed that the *Evening Star’s* claims about Seward’s purchases of icebergs etc. to stop their assailment on the purchase.


12. Ibid.


15. Ibid.

16. *Holt County Sentinel* (Oregon, MO), May 3, 1867.

17. Ibid.


23. Ibid.


26. Ibid.


Appendices

Appendix A

Appendix B2

Unknown.
“William the Glutton,” cartoon, Frank Leslie’s Weekly Newspaper (December 1867). I am indebted to the Granger Collection (granger.com) for the information about the location of the source.
Appendix C

Source: The London Times, April 1867.

Our policy is clear. Since we have no right whatever to protest an act entirely within the discretion of the Russian and Untied States Governments, let us not place ourselves in a false position of vain remonstrance. It is said that British Columbia almost cut off from the Pacific by the occupation by what ought to be a portion of its seaboard. The sufficient answer is that it was effectually cut off before, for America has only bought what belonged to Russia, and no Englishman ever dreamt that Russia would part with it to us. We are materially no worse off than before, while our moral right to our own possessions remains absolutely untouched.

Appendix D

Bibliography

Primary Sources


Charleston Daily News (Charleston, SC), 1868-1870.


Daily Phoenix (Columbia, SC), May 25, 1867.

Evening Star (Washington, D.C.), December 21, 1867.

Evening Telegraph (Philadelphia, PA), June 11, 1869.

Frank Leslie’s Weekly Newspaper (New York), February 1869.

Holt County Sentinel (Oregon, MO), May 1867.

London Times, April 1867.


New York Daily Tribune, April 11, 1867.


New York World, April 1, 1867.


**Secondary Sources**


Dozer, Donald Marquand. “Anti-Expansionism during the Johnson Administration.” *Pacific Historical Review* 12, no. 3 (Sept. 1943): 253-275. JSTOR.


Project Mind Control

Emma LaPorte and Darren Mei

1 Abstract

The original goal of this second semester Applied Science Research project was to make something move using “only our minds.” In order to achieve this seemingly impossible feat, we constructed an EEG (electroencephalography) device that uses electrodes to detect microvolts emitted from electrical impulses within the brain. Once amplified and filtered from background noise, these microvolts (now roughly millivolts) were to be sent to an Arduino microprocessor to ultimately create some sort of movement, most likely through a servo motor. The Arduino portion of this project is still a work in progress, but, as of May 2014, we were able to monitor brain waves (viewable via an oscilloscope) and observe a huge spike in amplitude when each of our test subjects tapped his or her foot. This accomplishment has made the project successful in our eyes.

2 Introduction

Researchers and engineers have created EEG-based headsets with the capability to perform a variety of final functions. Controlling a quadcopter or moving prosthetic limbs are examples. However, with our limited expertise in this area we hoped for our device to function in a more on/off manner. For example, we wanted to enable a servo to start spinning when a subject’s eyes opened, and stop spinning when the patients’ eyes closed. In order to get to this point, our biggest challenge was amplifying and filtering the voltage to a point that could be read by the Arduino/oscilloscope. This required research and the use of operational amplifiers and electronic filters. Overall, we learned a great deal about the process of amplifying/filtering voltage, and how the mind and its billions of neurons work.

This paper was written for Dr. James Dann’s Applied Science Research class in the spring of 2014.
3 History

Electroencephalography denotes the sensing of electrical activity along the scalp. Using electrodes that conduct minute voltages, scientists have been able to detect these voltages since the early 20th century. At first, physiologists like Richard Caton and Adolf Beck discovered these electrical phenomena in rabbits, dogs and monkeys. It was only in 1924 that German psychiatrist Hans Berger was successfully able to record the first human EEG. Berger made this recording during the surgery of a teenage boy, and in 1929 he reported on his findings, such as the alpha and beta waves that he detected. This discovery of electrical activity was a major breakthrough for the field of neuroscience, but due to World War II and a lack of recognition, Berger’s findings were not widely appreciated until the 1940’s. Nowadays, Berger is also credited with constructing the first electroencephalogram, or the actual device, which records the electrical activity along the scalp [2].

In the early stages of this technology, the main function of EEG devices was to discover what caused specific neurological diseases, such as epilepsy. More specifically, the difference between epileptic seizures and other symptoms like fainting, strokes, or migraines had not yet been identified. By using EEG devices, neurologists realized that patients with epilepsy suffered from an excessive amount of neuronal activity in the brain. Before they had a way of viewing brain activity, this fact was not known. From this point on, scientists realized the immense benefits of EEG devices, and in 1947 the American EEG Society was founded with the purpose of discovering more about this revolutionary technology.

Electroencephalographs, like any other piece of technology, have numerous advantages and disadvantages. The advantages of EEG devices are that they are relatively inexpensive to manufacture, they are mobile when compared to other brain scan devices, such as MRI’s, and they are not invasive, unlike other techniques which require surgery to monitor the brain. In addition, EEG devices are silent, and allow researchers to monitor the effects of auditory stimuli on a patient. The mobile function of an EEG device also allows users to move while being monitored,
making research on the effects of activity in the motor cortex on brain activity possible. However, the disadvantages of EEG devices are that it takes a lot of time for a proper EEG recording to be set up because it requires the precise placement of numerous electrodes on the scalp to achieve an accurate reading. Furthermore, since the electrodes are only connected to the scalp, brain activity further below the surface cannot be detected. For most researchers though, EEG devices are just as effective at recording brain activity over a certain period of time as other, more expensive devices are. In addition, the ability to have subjects move and interact instead of lying still makes the EEG device invaluable: it opens up many more avenues of research.

Since the introduction of electroencephalography, numerous fields have used EEG devices to discover more about the brain, for both trivial and medical uses. Medically, EEG devices have typically been used in diagnosing epilepsy, proving whether a patient is brain dead or not, and monitoring the brain activity of patients in comas. However, further research is being conducted by monitoring brain activity during sleep, and studying how different wave frequencies are sent through neuronal connections when subjects are relaxed compared to when they are focused. Much of the current research in electroencephalography is focused on aiding paralysis victims or amputees, because the calibration of certain EEG devices with motor functions could aid those who do not have full control over their bodies. In the future, prosthetic limbs could be controlled by amputees’ minds, and allow them to essentially regain a part of their body. Since brain activity associated with different kinetic actions, such as making a fist, generate distinct voltage readings; scientists could technically calibrate a person’s mind so that when he thinks of moving his arm a robotic arm could move in response. [1]

This field of study has been referred to as the study of a Brain-Computer Interface (BCI), and has been around since the mid-1970’s. Scientists in this field originally worked with animals such as monkeys, but in the 1990’s the first neuroprosthetic device was implanted in a human. Since then, peoples’ lives have benefited from this technology. One patient, Jens Naumann, a formerly blind man received a brain implant which
eventually allowed him to drive a car slowly around a small course. [3] Steps such as these in the BCI field are revolutionary, and are rapidly making what was once thought of as science fiction a reality.

Besides the medical usage of EEG devices, there are now commercially sold EEG devices that promise to improve sensory functions as well as track and monitor people’s everyday brain activity. Products such as the NeuroSky are currently selling portable EEG devices, targeted towards students, which are meant to monitor brain activity while the students work on homework. By taking these observations of the student’s brain activity and quantifying how focused, calm, and mentally stimulated they are, NeuroSky states that they can help improve students’ study habits and make them more efficient. [4]

Across the world, researchers are discovering new ways to harness the EEG device, and further develop the growing field. Recently, at the University of Minnesota, researchers constructed a connection between an EEG device and a quadcopter, which allowed the user to control a quadcopter using only his mind. Users were asked to think about making a fist with their right hand, which would allow the quadcopter to turn right. With such detailed instructions as these, it appears possible that in the near future, paralysis victims could control artificial limbs and possibly function independently. The brilliant ideas of researchers in the field of neuroscience, as well as the promising applications of EEG devices, have inspired this project, which is to construct an EEG device and allow a user to control a certain object using only his mind.

4 Theory

EEG devices are able to pick up voltages from electrical impulses within the body’s nervous system. Understanding these electrical impulses is key to understanding the science behind our project. These impulses are sent through neurons, which are made up of a dendrite cell and an axon (long fiber) leading to synaptic terminals and ultimately to another neuron. This image can be seen in Figure 1. Simply put, the charge of the inside of a typical cell body is negative due to its large organic
molecule proteins, while the surrounding salt solution has high concentration of positive sodium ions. What causes an electrical impulse is the movement of these positive sodium ions into the cell body, through protein channels, and ultimately through the attached axon sending an action potential through to the synaptic terminals. Following right behind this sodium ion trail are Potassium ions diffusing out of the cell body and axon. Parts of this action potential “leak out” at various increments along the axon, called Nodes of Ranvier, as it moves down the axon. This lost potential is immediately replaced as new of sodium ions move into the axon as the action potential continues to move across the length of the axon. In between these nodes there are myelin sheaths that act as insulation to keep the positively charged ions moving down the axon at a quicker pace. [5] Along with action potential, we will be picking up a little bit of postsynaptic potential. This potential is generated from neurotransmitters leaving the presynaptic terminal and moving across the synapse to a new neuron, which ultimately creates a transmembrane potential. [13] These two movements of ions or neurotransmitters create a potential, giving us (a very small) voltage to pick up, amplify, filter and send to an Arduino or oscilloscope.

Figure 1: A simplified neuron. [5]
After understanding the electrical impulses in the brain, it is essential to learn how the brain responds to external stimuli and how best to get a differentiated reading of this activity. By getting a clear alteration in brain activity, we are able to either monitor the change in voltage via oscilloscope or begin the on/off mechanism when the voltage is sent to an Arduino. We researched this in the scientific literature, but also met with Dr. Bruce Hill of Stanford University Medical Center. We learned that the brain produces frequencies mostly between 1-20Hz and voltages between 5-200μV. [7] Our hunch was that the most effective way to read brain activity would be to measure the alpha waves (frequencies range from around 8Hz to 12.5Hz and voltages range from 20-200μV). [7] Alpha waves are found during “wakeful relaxation.” [7] These make up the EEG pattern given off by an adult who is awake but relaxing with her eyes closed. To best measure alpha waves, two electrodes are placed on the skull, with an additional electrode as a ground, placed either on the earlobe or collar bone (both of which have limited muscle
tissue, which can limit interference with the wave pattern). As for the two electrodes placed on the skull, one is placed centrally on the top of the forehead, and the other on the bone protruding from the back of the skull, as seen in Figure 3. Alpha waves, also known as Berger’s waves after the founder of the EEG device, are the strongest signals sensed by an EEG. Additionally, and of utmost relevance to us, alpha waves alter significantly when patients open or close their eyes, as seen in Figure 3. Knowing that alpha waves abound during a state of relaxation, we made the connection that, in general, alpha waves increase when eyes are closed and decrease when the eyes open.

Figure 3: One of the two electrode placements enabling us to read alpha waves. As seen, the EEG will function regardless of the patient’s hair length.
Figure 4: A display of Emma’s alpha waves, while she was hooked up to an EEG. Although it is difficult to see from this photograph, the points of high amplitudes are while eyes are closed, while the points of low amplitudes are while eyes are open. These signals were gotten through use of a medical voltage amplifier, courtesy of Dr. Hill.

Figure 5: Dr. Hill’s overall set up at his lab. He uses a professional voltage amplification device.
Figure 6: This chart shows what is “normal” and what is not for alpha waves. These alpha waves are obviously much cleaner than those seen in Figure 4. This is because in practice there is a lot of interference and noise affecting the data, and we are never truly able to get the raw alpha waves.

Our research gave us insight into the range of alpha wave recordings. Figure 6 displays what is “normal” to see in alpha waves and what is not. If the waves are too small, it is often telling of a person who has trouble focusing and/or relaxing and more likely to have conditions like ADHD, Chronic Fatigue Syndrome, and Parkinson’s disease. [8] If alpha waves seem to be too large, this can be a signal of conditions like depression, “sluggishness” and thyroid problems. [8]

5 Final Method for Brain Activity Alteration

Our original idea was to have our test subjects close their eyes and rest for thirty seconds, and theoretically we would see increased amplitude due to the thriving alpha waves. We planned for this, and made sure we were not filtering out anything in the alpha frequency range. Our filter’s main purpose was to filter out background noise of around 60Hz and low frequencies below 3Hz. This left room for waves of other frequencies to be picked up.

After our circuitry was finalized, we began our testing process of identifying a simple method to effectively alter the voltage emitted from the
brain. Unfortunately, the eyes closed method did not work as effectively as we had hoped, or as well as it did while visiting Dr. Hill at his Stanford lab. His commercially used EEG was able to pick up voltages with higher specificity than our “home-made” EEG. Therefore, we went in search of a new method to increase the magnitude of the voltages emitted from the brain, based on some sort of movement or thought.

We tested a range of activities: listening to relaxing music as compared to loud, electronic music; reaching for objects; random movements; thinking about random movements. It was not until Ryan Hammarskjold randomly tapped his left leg while he was hooked up to the EEG that we hit on a new idea. When he tapped his leg, there was a huge spike in the amplitude. Since Ryan, we’ve seen the same increased voltage with every single test subject tapping or stomping this way.

There are a few theories behind why tapping one’s foot brings such a surge of voltage emissions. One is that the brain is sending a signal via its neuron network to the leg to make it move, and we are picking up that path of action potential. This is a possibility because we have noticed that oftentimes there will be a spike in the oscilloscope reading before a test subject has even set their foot down to stomp it. This led us to believe that it was not the impact of stomping the foot on the ground that caused the spike, but rather the brain signaling the foot to move. In other tests, the spike does not happen until after the foot has stomped. This told us that it was more likely the muscle movement altering the voltage. This is seen in Figure 4 on the edges of the voltage readings. There are points of extremely high voltage amplitudes when Emma closes and opens her eyes, an impact of the muscle movement, according to Dr. Hill. Although the theory behind our increased voltage from stomping the foot is unclear, the method is consistent and effective in altering the voltages emitted.
6 Design

6A. Operational Amplifiers

Before sending the voltage from the electrodes to an Arduino, the voltage must first be amplified and filtered. To amplify the voltage, operational amplifiers were utilized in series to boost the voltage output from the electrodes by approximately 1000 times. Since the expected output from the electrodes is between 10 to 50 microvolts (10^{-6}V), and an Arduino can only detect values in the millivolt range (10^{-3}V), an amplification of approximately 1000 was necessary. For this circuit, two Texas Instrument Dual General-Purpose Amplifiers were utilized. Since these were dual operational amplifiers, the circuit ran through three out of the four possible op amps provided by these two devices. [9]

An operational amplifier is a device that takes the voltage difference between two inputs and magnifies that difference by a certain factor (referred to as the gain). These two inputs are referred to as either the inverting input (represented by the – sign), or the non-inverting input (represented by the + sign). The inverting input is for negative voltages, and the non-inverting input is for positive voltages. Typically, one of these two inputs is grounded, because that allows for the voltage from one input to be amplified instead of the difference between the voltages from two inputs. For an ideal op amp (see Figure 7), the gain is either infinity or negative infinity depending on whether an inverting or non-inverting input has been grounded. However, an output of infinity or negative infinity is not possible, and for actual op amps, a system of feedback is utilized to control what the output of the op amp is. In Figure 8, a resistor is connected from the output of the op amp to the input into the op amp. This makes the output from the op amp become the input into the op amp, creating a cycle which makes the difference between the two inputs decrease to zero. Since the op amp wants the difference between the inputs to be zero, when the feedback cycle makes the difference go to zero, the op amp outputs the amount of voltage that was necessary to correct for the difference. For example, if the inverting
input was grounded, 1V was sent to the non-inverting input of the op amp, and the resistors were equal to each other, the op amp would have to output 1V to account for the difference between the inverting and non-inverting inputs. [10]

![Figure 7](image1)

Figure 7: An ideal operational amplifier. Notice the lack of a resistor connecting A to C, which in a normal op amp would create the feedback. [10]

![Figure 8](image2)

Figure 8: A circuit diagram of an example inverting operational amplifier. The non-inverting input (represented by the + sign) is going to ground, while the voltage is seen going into the inverting input. This circuit diagram also has a resistor connecting the output voltage back into the inverting input. [14]

To actually amplify the voltage difference, resistors are put into the op amp circuit, which will change the amount of voltage coming out of the op amp. As seen in Figure 8, there is a resistor between the voltage source and the inverting input to the op amp. This resistor is commonly referred to as $R_{\text{in}}$. The resistor that connects the output of the op amp to the inverting input is referred to as $R_f$. The relationship between $R_f$ and $R_{\text{in}}$ is what determines the amount of gain that an op amp provides, because $\frac{R_f}{R_{\text{in}}}$ = $A$ (gain). This relationship exists because the current stays constant through the circuit, so by increasing $R_f$, the voltage output of the op amp increases because of $V = IR$. This multiplication...
factor of the op amp is what makes it so valuable, because it can drastically amplify small voltages.

An issue with operational amplifiers is that there is sometimes saturation, which is when the op amp's gain is too large, so its output is capped at a certain voltage. This can be seen in Figure 9, where the full range of voltages is not met because of saturation. When saturation occurs, $R_f$ must be decreased so that the entire range of voltages can be observed. The reason that saturation occurs is that the output voltage is too close to the power supply voltage. By being so close to the power supply voltage, the op amp is reaching its maximum output, causing it to cut off and "saturate". To solve this issue, either $R_f$ can be lowered, or the voltage input into the op amp can be lowered. Since the input into the op amp is going to be extremely low (in microvolts), a voltage divider was utilized (see Figure 13) to simulate the small voltage input. By decreasing the voltage output from the voltage divider, the voltage input into the op amp was lowered, and prevented the op amp from saturating. Even after the voltage divider, approximately 3.50mV was still being inputted into the op amp, so when the electrodes are used instead as the inputs, saturation will not be an issue.

Figure 9: This graph displays the saturation effect on voltage output. As the voltage increases past a certain point, the output plateaus, because it is approaching the voltage coming from the power supply. The dark and light gray lines show what will appear on the oscilloscope when saturation occurs, while the black line shows the desired output reading. [15]
As seen in the circuit diagrams below (Figures 10 and 11), three operational amplifiers were used, with the non-inverting input being grounded in each. The resistor values were found through trial and error, and trying to reach the maximum gain before saturation occurred. The output from this circuit is going to lead to an Arduino, which can read in the amplified voltages and adjust its function according to the different voltage readings it receives.

![Figure 10: This is the Operational Amplifier Circuit generating an amplification factor of approximately 1,000. This will make the microvolts picked up from the electrodes detectable for both the oscilloscope (for testing purposes) and Arduino (for the final product).](image1)

![Figure 11: This is the same circuit as seen in Figure 7, but detailed with the specific pins of each op amp and where they lead. Op Amp 1 and Op Amp 2 are both Texas Instruments Dual General-Purpose Amplifiers. Both op amps were utilized in Op Amp 1, and one op amp was utilized in Op Amp 2.](image2)
Figure 12: This is a photo of the breadboard containing the operational amplifiers. The leftmost column goes to ground, the second leftmost column connects to $+15\text{V}$, and the second rightmost column connects to $-15\text{V}$. The first op amp is shown on the top left, the second op amp is on the top right, and the third op amp is on the bottom right. The white wire on the bottom right is the output, which will connect to the voltage sensor on the Arduino.

Figure 13: This is a diagram of the voltage divider used to simulate the low voltages inputs from the electrodes. The output of this voltage divider was approximately $3.5\text{mV}$. 
Figure 14: This is a photo of the breadboard containing the voltage divider. The wire on the bottom right is the input from the power supply, the wire on the top left goes to ground, and the wire below that is the output which goes to the op amp circuit (see Figure 9).

6B. Filters

When voltages are amplified over one thousand times, a lot of unwanted signals are also picked up. At such small voltages, these extraneous signals can come from electronics that are in the vicinity, and make it difficult to see the desired signal. To get rid of the noise from these signals, filters are used to only allow signals of a certain frequency to pass. There are different types of filters, each of which has a different function to isolate frequencies. The filters that are being considered for this project are either a low-pass filter or a band-pass filter. A low-pass filter only passes through signals below a certain frequency. This filter can be used to get rid of signals from surrounding electronics, and allow for only the signal from the electrodes to pass through, which is expected to have a frequency around 8-15Hz (alpha wave range). The band-pass filter contains a low-pass and a high-pass filter, and instead of having ei-
ther a maximum or a minimum frequency cutoff, it contains both. This could allow us to ideally eliminate frequencies below 8Hz and above 15Hz. Both of these filters will be tested for optimal performance.

Another decision to consider when designing the filter circuit is whether to use an active filter or a passive filter. A passive filter does not require a power supply, and can handle large currents and high voltages. However, passive filters require the use of inductors as well as a more complicated circuit, which makes it difficult to set up. An active filter requires a power supply, because it uses op amps, and is not as reliable as a passive filter. Active filters are easier to design, though, and do not require the use of inductors. Since an op amp is being used already to amplify the voltage, an active filter was chosen for this project, because it also is smaller in size and weight, making the setup easier to manage. [11] We consulted a previous student’s project, one by Ben Adler, because of his work with active band-pass and low-pass filters and his circuit diagrams (see Figure 15).

**Figure 15:** Ben Adler’s circuit diagram for his speaker system project. This diagram is being consulted for the initial construction of an active low-pass filter circuit.
For the final circuit, an active low-pass filter and an active notch filter were used, although the notch filter performed the function of a high-pass filter. The low-pass filter demonstrated a steady drop in amplitude for frequencies above 20 Hz, and the notch/high-pass filter demonstrated a rapid drop in amplitude for frequencies lower than 1 Hz. Below are the final filter circuit diagrams along with a picture of the circuit on a breadboard.

**Figure 17:** The circuit diagram for the low-pass filter, which filters out signals that have a frequency higher than 20Hz. This is not a sudden cutoff though, and instead is a steady lowering of amplitude until the signal coming through is much less noticeable.
Figure 18: The actual breadboard containing the circuit. Two 9V batteries, with one supplying the positive voltage, and another supplying the negative voltage, power the op amp in this circuit.

Figure 19: The circuit diagram for the notch filter, which instead acted as a high pass filter. This circuit showed a rapid decrease in amplitude for frequencies lower than 1Hz, which helped to filter out some of the DC signals that had been amplified by the amplification circuit.

7 The Actual Product

In our initial plan, we sought to create a headset that could easily be put on and used. After working with the electrodes and the conductive paste, we are realizing that this might be unnecessary. The conductive
gel acts as an effective adhesive that holds the electrodes to the person’s scalp. Plus, dispensing with a headband leads to more flexibility and accuracy of proper placement of electrodes (not just where the headband lands).

Our product finally seemed to come together. We decided it would include electrodes connected to our amplifying and filtering circuits, which were connected to an Arduino, which was connected to something that would “move” or “light up.” These medically used electrodes, like all other electrodes, are conductors that pass electrical current from one medium to another. In this case, the electrodes are picking up the voltages from electrical impulses within the brain and passing them through the connecting wire toward our circuit. [16] These electrodes are put on in pairs, in order to measure the difference between the two metal plates. For most EEG devices, alpha waves (8-12Hz) are the easiest waves to pick up, but for the device made in this project, the signals picked up seemed more in tune with the nervous system and muscular contractions. The signals were typically picked up when the subject would tap his or her foot on the ground, or in some cases raises a leg. Also, some subjects would experience larger changes in amplitude based on their actions than others, suggesting that either some people have different levels of brain activity for specific actions, or that the placement of the electrodes varied enough to affect the output of the circuit. Initially, the user of our device would have an electrode placed on the middle of the forehead, the back of the head where the skull bone juts out, and on the skin above the collarbone. However, as testing continued, it was observed that the electrode connected to the back of the head was not required, so only the electrodes on the forehead and collarbone are currently being used by subjects. The forehead electrode is acting as the “voltage in” for the circuit, while the collarbone electrode is connected to the ground line of the circuit. The reason the ground electrode is placed on the collarbone is that the collarbone has low muscle density, so there is no such tissue to interfere with the waves captured.
When we first planned it, the final stage of this project was to include an output from the circuit into an Arduino, which would allow a user to actually “control something with his or her mind.” After we read the output from the circuit on an oscilloscope, however, a major problem arose that made the Arduino connection impossible. Although there is a clear difference in the AC signal when a subject moves a leg or otherwise “activates” the circuit, there is still an underlying DC signal of approximately 8.2V coming out of the circuit. Despite the fact that most of the DC signal was filtered out using the notch/high-pass filter, there was still approximately 8mV after the filters, which was then amplified by the op amp, causing for such a large signal at the end of the circuit. This 8.2V DC signal was a major problem for the Arduino input, because the Arduino Mega can only read in at most 5V. Since the signal coming out of the circuit was too high, the Arduino could not read in the circuit’s output properly.

To address this issue, we researched numerous methods of filtering out the DC component of a composite signal, and eventually we settled on two options. We found that there is a device capable of filtering out the DC component of a composite signal, called an AC coupling, which can be found in oscilloscopes. With this knowledge, we took apart an old oscilloscope (see Figure 20) so that the AC coupling could be disconnected. Unfortunately the circuitry of the oscilloscope was interconnected to such an extreme degree that we were unable to isolate and utilize just the AC coupling component of the oscilloscope.
Figure 20: This is a photo of the oscilloscope that was taken apart for the AC coupling. The AC coupling was in the bottom left of this photo, on the bottom rack of this oscilloscope. As is seen in the photo, there are many wire connections in this photo, and due to the amount of connections, it was determined that it would be too difficult to isolate and utilize the AC coupling in this oscilloscope.

Our last attempt to fix the DC current issue involved creating an AC coupling circuit by hand. Below is a picture of the attempt, which unfortunately, was not successful. However, if this project were to be extended, the AC coupling circuit would be attempted again, because if the DC current issue were solved then the Arduino would ideally be able to read in the output from this circuit.
Figure 21: This is a photo of the attempted AC coupling circuit. While the DC current was reduced, the AC signal did not show any change when a subject would move a leg or do actions that would normally generate a spike on the oscilloscope.

8 Results

The results for this project were recorded via photographs taken of the oscilloscope screen. This gave us voltage as a function of time, as picked up by the electrodes placed on the scalp. As you will see in the following photographs of oscilloscope readings, the brain activity (or voltage) captured changes as a result of the act of tapping or stomping a foot.
Figure 22: This shows a reading given by the oscilloscope of a test subject’s brainwaves in an initial state. Our initial state would be to have the test subject sit still, doing absolutely nothing. The graph of an oscilloscope gives voltage as a function of time.

Figure 23: This shows the altered brain activity when a test subject stomped a foot on the ground. As you can see, the voltage range increased significantly from this movement.
Figure 24: This oscilloscope reading was taken while a test subject tapped a foot repeatedly. Again, the voltage going through to the oscilloscope has changed dramatically from the initial, flat reading.

9 The Next Step

At this point, we successfully amplified the microvolts to readable levels, filtered out all background noise, and read brainwaves. We nailed down a method that consistently sparks a spike in the voltages emitted from the average test subject’s brain. This alone was a huge accomplishment and something we were not sure would be able to do when we started this project. Although the Arduino proved to be a little more difficult than originally imagined, we will troubleshoot and continue to work on this aspect of the project. We have had a range of ideas about what we could do next with the knowledge we gained from this project. We would love the chance to try using the setup to turn on a servo motor, to make something move, to light up a light bulb above a subject’s head, etc. This project has opened the door for us to the many amazing medical applications of engineering. And, on top of it, Project Mind Control was a blast! ●
Figure 25: Kevin Jacques, a test subject, demonstrating his intelligence with the device.

Works Cited


Singing Tesla Coil: Building a Musically Modulated Lightning Machine

Sam Redmond

1 Abstract

A musically modulated Tesla coil was developed and tested but never made fully functional. Music data is sent from an input MIDI stream, either from preexisting MIDI files (transferred from a computer) or from a MIDI-enabled piano keyboard to a microcontroller. The microcontroller then processes incoming MIDI signals and translates them into a series of electrical pulses. Software was created for the Arduino Due (a microcontroller) that allows up to nine concurrent notes to be played. The output pulses control whether the Tesla coil sparks or not. The variable frequency of sparking produces the perception of sound.

2 Introduction

Tesla coils are inherently fascinating. They mix physics, danger, and lightning. What’s not to like? Adding music to the mix not only makes the project more challenging but also adds novelty. While many hobbyists have made Tesla coils, only oneTesla, ArcAttack, GeekGroup, and a few other YouTubers have made singing Tesla coils. [1] [2] [3] Only oneTesla and ArcAttack have controlled a Tesla coil with an instrument; even so, they have only been able to play two notes simultaneously. Our system supports up to nine concurrent notes. This project doesn't have much relevance to the wider world. At no point will society desperately need MIDI-controlled lightning machines in the same way that it will need sustainable energy sources or a more efficient light bulb. However, entertainment is always in demand, and nothing draws a crowd quite like Pachelbel’s Canon played on the Tesla coil.

This paper was written for Dr. James Dann’s Advanced Science Research class in the spring of 2014.
2.1 My Interest

Personally, I can’t imagine a better project. It involves a significant amount of programming and electrical engineering, but also has some elements of mechanical engineering. I love music too, and MIDI blends music and electronics. I also enjoy the danger associated with these extremely high voltages.

3 History

3.1 Nikola Tesla

Nikola Tesla, the inventor of the Tesla coil, lived a dramatic life. Born in 1856 in what is now Croatia (then Serbia), Tesla studied math and physics at the University of Prague. Legend holds that Tesla, the brilliant student, came up with the idea for a brushless AC motor while on a walk, sketching the schematic in the sand with a stick.

In 1884, Nikola Tesla moved to the United States and soon found a job as an electrical engineer in Thomas Edison's headquarters. After seeing how smart Tesla was, Edison offered Tesla $50,000—equivalent to roughly $1.1 million today—to redesign Edison’s inefficient DC power generators. After Tesla worked tirelessly for months to find a solution, Edison laughed off his previous offer, refused to pay Tesla the money, and chided: “Tesla, you don’t understand our American humor.” Tesla left Edison’s company.

On his own, Tesla attempted to create the Tesla Electric Light Company, but failed and resorted to digging ditches for $2 a day for a brief period. Later, Tesla found a few investors for his research on alternating current. In 1887 and 1888, he was awarded over 40 patents for his work on AC, and spoke at a conference, attracting the attention of George Westinghouse. Westinghouse was a huge proponent of the AC system, and had recently set up the first AC power system in Boston. Westinghouse hired Tesla, bought his patents, and gave him a research lab. However, Tesla's and Westinghouse's AC was a direct competitor to Edison's DC. Therefore, the greedy, power-hungry
Edison began to publicly denounce the AC system. In 1889, a convict was killed by electrocution using AC current, overseen by Edison. Edison electrocuted cats, dogs, calves, horses, and an orangutan with AC current to show the public just how dangerous AC could be. Perhaps Edison’s greatest act of cruelty came in 1903, when he electrocuted Topsy the 6-ton circus elephant with 6000 volts AC in front of a group of 1500 spectators. [4] [5] (Warning: the videos are gruesome with regards to animal cruelty.)

Ultimately, though, AC power won out. For one, lightbulbs that ran on AC current were bright white, compared to Edison’s dull yellow DC bulbs. The biggest reason that AC was preferred over DC involves the transfer of electricity. AC power lines can run at much higher voltages than DC power lines, which allows for a lower current to transfer the same amount of power \((P = IV)\). The lower current of AC power means that less power is dissipated in the transfer cables \((since P = I^2R)\), so cables can be longer. Whereas DC power required substations every 2 or so miles, AC power lines could run for hundreds of miles with negligible power losses.

While Westinghouse went off to supply the nation with AC power, Tesla continued to invent. In 1893, he demonstrated the AC system at the World Columbian Exposition in Chicago, and in 1895 he built a hydroelectric power plant at Niagara Falls. Tesla lost many notes and machines when his lab burned down in 1895 and began moving around the country, staying in Colorado Springs before moving to New York. Tesla got funding from J.P. Morgan to build a worldwide communications network with a central hub at Tesla’s Wardenclyffe lab on Long Island. However, Morgan pulled the funding, and Wardenclyffe was destroyed.

Tesla’s last years were spent in the New Yorker Hotel, feeding (and ‘communicating with’) the city pigeons. He became obsessed with the number three and abhorred germs, causing historians to question his mental health. Nikola Tesla died on January 7, 1943, alone and with significant debts. His work on dynamos, electric motors, radar, X-rays, and magnetic fields helped to inspire the next generation of
inventors. Modern-day radio transmission relies on principles that Tesla discovered. Even though Guglielmo Marconi is often considered the ‘father of the radio,’ Tesla’s research predates Marconi’s and in 1943, the Supreme Court voided four of Marconi’s patents, giving due credit to Tesla. Perhaps most importantly, Tesla’s AC power system is the worldwide standard today. [6] [7] [8]

3.2 The Tesla Coil

In 1891, Nikola Tesla invented one of his most famous devices, the Tesla coil. To be fair, electrical coils existed before Nikola Tesla. Ruhmkorff coils, named after Heinrich Ruhmkorff and designed by Nicholas Callan in 1836, converted low voltage DC to high voltage DC spikes. [9] Tesla began working with these transformers, but soon switched to devices of his own invention. Tesla’s capacitors consisted of metal plates immersed in oil, and his spark gap used two metal balls and an air jet to clear ionized air, creating more abrupt discharges. The secondary coil was comprised of two inductors, one to couple the fields and one ‘resonator.’ Later, Tesla developed higher voltage power transformers, and used a rotating spark gap to short the LC circuit. Voltage gain was significantly increased when Tesla began loosely, rather than tightly, coupling the two inductors, using air as the core rather than metal. Tesla’s patent was for a ‘high-voltage, air-core, self-regenerative resonant transformer that generates very high voltages at high frequency.’

Modern Tesla coils have not changed much from Tesla’s ultimate design. Modern designers have simplified the circuit, removing the helical resonator that previously was connected in series with the secondary coil. Since Tesla’s original goal was power transmission, his top conductors often had a large radius of curvature to minimize any electrical losses from corona effects or discharges to the secondary coil or surrounding objects. Modern builders emphasize long sparks more than efficient power transmission, so modern Tesla coils tend to use toroidal top conductors rather than Tesla’s original spherical conductors. [10]
3.3 Modern Musical Tesla Coils

Musical Tesla coils were first seen in public demonstrations in March 2006 by Joe DiPrima and Oliver Greaves. In 2007, DiPrima renamed his group ArcAttack, a musical performance group, which continues to showcase musical Tesla coils today. [11] Steve Ward and Je Larson were also early popularizers of the singing Tesla coil, alternatively called a Zeusaphone (a play on words on the tuba-like instrument, the Sousaphone, using Zeus, the Greek god of lightning) or a Thoremin (a play on words combining ‘theremin,’ the electronic musical instrument, with Thor, the Norse god of thunder. [12] Ward’s original blog post discusses the physics and engineering challenges of constructing a musical Tesla coil. [13]

Since then a number of alternate musical Tesla coil groups have sprung up, most notably oneTesla, a group out of MIT that sells musical Tesla coil kits. The company was created after a wildly successful Kickstarter campaign. [14] The oneTesla kit, however, only allows for a small Tesla coil with support for only two concurrent notes. Nevertheless, communities such as those at oneTesla and the high voltage enthusiasts at 4HV have helped boost the popularity of musical Tesla coils. [15]

4 Theory

This section is a theoretical discussion of physics concepts from electromagnetism, Tesla coils, Arduino features, MIDI, and miscellaneous circuit elements.

4.1 Physics Background

4.1.1 Maxwell’s Equations

As with all electromagnetic systems, we begin with Maxwell’s equations, the fundamental equations that unify electricity and magnetism.

In order, they are Gauss’s Law (1 and 5), Gauss’s Law for Magnetism (2 and 6), Faraday’s Law of Induction (3 and 7), and Ampère’s Circuital
Law (4 and 8) with a correction by Maxwell, shown first in their integral form and second in their differential form.

\[
\iint_{\partial \Omega} \vec{E} \cdot d\vec{S} = \frac{1}{\varepsilon_0} \iiint_{\Omega} \rho \, d\nu = \frac{q_{\text{encl}}}{\varepsilon_0} \tag{1}
\]

\[
\iint_{\partial \Omega} \vec{B} \cdot d\vec{S} = 0 \tag{2}
\]

\[
\oint_{\partial \Sigma} \vec{E} \cdot d\ell = -\frac{d}{dt} \iint_{\Sigma} \vec{B} \cdot d\vec{S} \tag{3}
\]

\[
\oint_{\partial \Sigma} \vec{B} \cdot d\ell = \mu_0 \iiint_{\Sigma} \left( \vec{J} + \varepsilon_0 \frac{\partial \vec{E}}{\partial t} \right) \cdot d\vec{S} \tag{4}
\]

\[
\nabla \cdot \vec{E} = \frac{\rho}{\varepsilon_0} \tag{5}
\]

\[
\nabla \cdot \vec{B} = 0 \tag{6}
\]

\[
\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \tag{7}
\]

\[
\nabla \times \vec{B} = \mu_0 \left( \vec{J} + \varepsilon_0 \frac{\partial \vec{E}}{\partial t} \right) \tag{8}
\]

In these equations, \( \mu_0 \) is the permeability of free space and \( \varepsilon_0 \) is the permittivity of free space. Both are universal constants

\[
\left( \mu_0 \equiv \frac{10^{-7}}{4\pi} \frac{\text{V} \cdot \text{s}}{\text{A} \cdot \text{m}} \text{ and } \varepsilon_0 \approx 8.8542 \times 10^{-12} \frac{\text{F}}{\text{m}} \right) \text{ related by } \quad c = \frac{1}{\sqrt{\varepsilon_0 \mu_0}}.
\]

where \( c \), the speed of light in free space, is defined to be \( 299,792,458 \ \frac{\text{m}}{\text{s}} \).

The concepts of relative permeability and relative permittivity in a material can be used to define the speed of light in a material.
\( \mathbf{E} \) is the electric field at a given point in space and \( \mathbf{B} \) is the magnetic field at a given point in space.

\( \nabla \) is a mathematical operator on functions defined on \( \mathbb{R}^n \). In 3D, this operator can be thought of as a 3-vector of partial derivatives,

\[
\nabla = \left\{ \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right\}
\]

For vector-valued functions \( f = P(x,y,z), Q(x,y,z), R(x,y,z) : \mathbb{R}^3 \rightarrow \mathbb{R}^3 \)

where \( P, Q, \) and \( R \) are real-valued functions from \( \mathbb{R}^3 \) to \( \mathbb{R} \),

\[
div f = \nabla \cdot f = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}
\]

is the divergence of \( f \) and \( \nabla \times f = \left( \frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}, \frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}, \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) \)

is the curl of \( f \). Note that divergence is real-valued and curl is vector-valued. Divergence can be thought of as flux per unit volume at a point. The curl is usually only applied to 3-dimensional functions and measures rotation at a point. The vector-valued functions we're dealing with are \( \mathbf{E} : \mathbb{R}^3 \rightarrow \mathbb{R}^3 \) and \( \mathbf{B} : \mathbb{R}^3 \rightarrow \mathbb{R}^3 \).

\( \Omega \) is a closed volume with bounding surface \( \partial \Omega \) and locally planar differential surface \( d\hat{S} \), directed perpendicularly out of the surface and with magnitude equal to the area of the plane. \( \Sigma \) is an open surface with bounding contour \( \partial \Sigma \) and locally linear differential line \( d\ell \) parallel to \( \partial \Sigma \).

The quantity \( \rho \) is the charge density at a point in space, in units of \( \frac{c}{m^3} \), related to total charge enclosed in a surface by \( Q_{\text{enc}} = \oint_{\partial \Omega} \rho d\mathbf{V} \cdot \hat{J} \)

is current density, in units of \( \frac{A}{m^3} \), related to net current through a surface by \( I_{\text{enc}} = \oint_{\partial \Sigma} \hat{J} \cdot d\hat{S} \)

Why do we bother with both forms of the equations? The integral forms describe electromagnetic behavior in a global region of space, whereas the differential forms describe local electromagnetic behavior at a point.
We’ll focus on Faraday’s Law of Induction and Ampère’s Circuit Law. In English, Faraday’s law states that a time-varying magnetic field will induce an electric field, and Ampère’s law states that a magnetic field is induced not only by an electrical current, but also by a time-varying electric field.

4.1.2 More on Ampère’s Law

We’ll ignore Maxwell’s correction for now and focus on the main component of Ampère’s law that approximately posits that a current will induce a magnetic field. (See Figure 1, bottom left.) For convenience, we’ll rewrite Ampère’s Law as:

\[ \oint_{\partial \Sigma} \vec{B} \cdot d\ell = \mu_0 I_{\text{encl}} \]  

(9)

where the variables are the same as before, with the exception that \( I_{\text{encl}} \) refers to the total current enclosed by \( \Sigma \).

Consider an infinitely long, straight wire oriented in the Z direction, carrying a current of \( I \). (Note: If the wire were finitely long, we would have to apply Maxwell’s correction to Ampère’s Law to obtain a rigorous result.) Let a distance \( r \) be given. Define \( \Sigma \) as the disk \( x^2 + y^2 < r^2 \). Then \( \partial \Sigma \) is the circle \( x^2 + y^2 = r^2 \). The current enclosed by \( \Sigma \) is simply \( I \) (since the wire pierces our Ampèrian surface), so the right side of Equation 9 simplifies to \( \mu_0 I \). By symmetry, the magnitude of the magnetic field \( \| \vec{B} \| \) is the same at all points on \( \partial \Sigma \), and the angle \( \theta \) between \( \vec{B} \) and \( d\ell \) is constant, so we find:

\[ \oint_{\partial \Sigma} \vec{B} \cdot d\ell = \| \vec{B} \| \cos \theta \oint_{\partial \Sigma} d\ell = (\| \vec{B} \| \cos \theta)(2\pi r) = \mu_0 I_{\text{encl}} \]

\[ \| \vec{B} \| \cos \theta = \frac{\mu_0 I}{2\pi r} \]

From Gauss’s Law for Magnetism, \( \vec{B} \) is a conservative field, so \( \vec{B} \) must be tangent to our path of integration (so that, when dotted with the normal, the integrand becomes zero). Therefore, \( \theta = 0 \) and \( \cos \theta = 1 \), so we get our first result: At a distance \( r \) from a line of current, the magnitude of the magnetic field is \( \frac{\mu_0 I}{2\pi r} \). Its direction is arbitrarily...
designated (definitionally) by the right hand rule: with the thumb of the right hand aligned with the direction of the current, the magnetic field follows the direction that the fingers naturally curl.

Our next goal is to determine the magnetic field inside a solenoid. Consider an ideal infinitely long solenoid: a helical coil of wire with turn density \( n \) (units of \( \text{turns/m} \)), current \( I \), wrapped around an (infinitely long) imaginary cylinder of radius \( r \) (see Figure 1, top left). Without loss of generality, let the axis of the cylinder be the z-axis (given by \( \hat{k} = (0,0,1) \)), and let the current be flowing in the positive z direction (parallel to \( \hat{k} \)). By symmetry and by applying the right hand rule repeatedly, we argue that the magnetic field on the interior of the solenoid is parallel to \( \hat{k} \) and that the magnetic field on the outside of the solenoid is parallel to \(-\hat{k}\).

![Magnetic field inside a long solenoid.](image1)

![Magnetic field from a long straight wire.](image2)

![Magnetic field inside a toroidal coil.](image3)

![Magnetic field inside a conductor.](image4)

**Figure 1:** Visualizations of the magnetic field created by a current-carrying wire in common configurations. [16]

Consider a positively-oriented Ampèrian square ABCD in the x-z plane height \( h \) and width \( w \). Let AB and CD be parallel to the x-axis, let BC be parallel to the z-axis and outside the cylinder, and let DA be parallel to the z-axis and inside the cylinder. Since AB and CD are perpendicular to the magnetic field both inside and outside of the
solenoid, their contributions to the loop integral drop out, since the dot product of $\mathbf{B}$ and the tangent line is zero. The contribution from BC is also zero, since the magnetic field outside of this ideal solenoid is zero. This is because all magnetic field lines must form complete loops (Gauss’s Law for Magnetism), and while the ratio of volume outside the cylinder to volume inside the cylinder increases without bound, the magnetic field lines are effectively spread through an infinite region, and are thus ‘diluted’ to zero strength. Therefore, the only nonzero component of the loop integral occurs on the interior segment (DA). Additionally, the vector dot product $\mathbf{B} \cdot d\mathbf{l}$ becomes the algebraic product $B \, dl$ along this segment because the tangent line and magnetic field lines are parallel. The net current $I_{encl}$ enclosed by this rectangular region is given by $hnI$, since there are $hn$ points where a wire with current $I$ pierces the region.

Applying Ampère’s Law, we find that:

$$\oint_{ABCD} \mathbf{B} \cdot d\mathbf{l} = \int_{D}^{A} B \, dl = B \int_{D}^{A} dl = Bh = \mu_0 hnI = \mu_0 I_{encl}$$

The direction is given by a variation on the right hand rule: with the fingers curling in the direction of current flow in the solenoid, the thumb points along the direction of the magnetic field.

4.1.3 More on Faraday’s Law

Faraday’s Law is all about how a changing magnetic field can induce an electric field. Again, we rewrite Faraday’s Law to intuitively understand it better:

$$\mathcal{E} = -\frac{d\Phi_B}{dt}$$

where $\mathcal{E}$ is the induced electromotive force (emf) and $\Phi_B$ is the magnetic flux through some Gaussian surface. If a magnetic field is established inside a solenoid by an external source, the solenoid will experience an induced voltage. This principle is key in transferring energy from the primary circuit to the secondary circuit in a Tesla coil.
The negative sign in front of \(-\frac{d\phi_B}{dt}\) means that the induced voltage opposes the charge in magnetic field. This principle is called Lenz’s Law, and means that an inductor opposes the flow of current through it (since current would establish a magnetic field).

4.1.4 Inductors

Inductors have a property known as inductance, commonly given the symbol \(L\), which characterizes this opposition to current. Mathematically,

\[\mathcal{E} = -L \frac{dl}{dt}\]

In this case, \(\mathcal{E}\) is often referred to as the backemf.

4.1.5 LC Circuits

Consider the undriven LC circuit shown in Figure 2. The capacitor has capacitance \(C\), and the inductor has inductance \(L\). Consider what happens when the capacitor is fully charged. At first, the inductor poses no backemf, so current discharges through the capacitor. However, as the magnitude of current flowing through the inductor increases, the magnetic field through the inductor increases as well, so the inductor experiences a backemf opposite to the direction of current flow. This occurs until the flow of current in one direction stops completely. However, at this point, charge has built up on the opposite plate of
the capacitor, so the cycle begins again in reverse. We can calculate the frequency of this oscillation using what we know about circuits, capacitors, and inductors.

Mathematically, we know from Ampère’s Law (Kirkhoff’s Voltage Rule) and conservation of charge (Kirkhoff’s Current Rule) that $V_C + V_L = 0$ and $i_C = i_L$. Additionally, we know that

$$V_C(t) = L \frac{di}{dt}$$

and

$$Q = V_C C \Rightarrow \frac{dQ}{dt} = i_C(t) = C \frac{dv_C}{dt}$$

Rearranging, we have:

$$\frac{d^2i(t)}{dt^2} + \frac{1}{LC} i(t) = 0$$

Define $\omega_0 = \frac{1}{\sqrt{LC}}$. The solution to this differential equation yields the result:

$$i(t) = i_0 \sin(\omega t + \phi)$$

If we let $\phi$, the phase shift, be zero, we see that this LC circuit oscillates in a sinusoidal pattern, much like the simple harmonic oscillators we know and love. Specifically, the LC circuit oscillates with frequency $f = \frac{1}{2\pi\sqrt{LC}}$ Hz. In every cycle, charge flows from one plate of the capacitor, through the inductor, to the other plate and back. In fact, the typical resonance conditions hold; if there exists a driving voltage that oscillates at the same frequency and is in phase with the oscillations in the LC circuit, then the amplitude of the oscillation monotonically increases without bound. Of course, in real life, wires have resistance, so there is some damping factor. However, with the Tesla coil, the rate of increase of amplitude far outweighs the damping effects.

4.2 Tesla Coil

The main concepts behind all Tesla coils are the same: an oscillating primary LC circuit induces resonance in a secondary LC circuit, building up voltage on the secondary’s capacitor until it breaks down.
We’ll first focus on the theory behind the spark gap Tesla coil, and then describe how our Tesla coil’s design slightly differs.

4.2.1 Spark Gap

The basic circuit for a spark gap Tesla coil is shown in Figure 3a. The spark gap, primary inductor, and capacitor are all connected in series. Charge builds up on the capacitor from the AC power source (usually sent through a transformer in order to step up the voltage), until the voltage across the spark gap is high enough to breakdown and ionize the air between the leads of the spark gap. After all, the spark gap is a capacitor in disguise, and therefore has some breakdown voltage. Ionized air is a great conductor (compared to unionized air, which is a great insulator). The capacitor can then discharge through the spark gap until the air deionizes. The ionization voltage threshold is higher than the deionization voltage. As the capacitor is short-circuiting through the spark gap, it is also discharging through the inductor, forming an oscillating LC circuit. The frequency of oscillation of a Tesla coil’s primary circuit is so high (MHz range) that the AC source (at 60Hz) can be thought of as DC, and thus does not affect the frequency of the resonant circuit.

While the spark gap is ‘closed’, the primary Tesla coil circuit acts exactly as an LC circuit. We know that an oscillating LC circuit will induce alternating magnetic fields in the interior of its inductor. In a Tesla coil, another inductor (the secondary) is concentric with the primary, but usually has a much high turn density. From Faraday’s Law of Induction, this means that a voltage (\( \mathcal{E} \): electromotive force) is induced in the wires of the secondary coil. This induced voltage will charge the secondary circuit’s capacitor (really, the top toroid is just one plate of a capacitor, where the other plate is the physical ground), and then the secondary coil will begin displaying resonant LC behavior, albeit initially at a lower energy level. Each time the primary oscillates, it induces some voltage in the secondary. Essentially, the current in the primary is driving oscillations in the secondary.
Consider what happens if the angular frequency of each circuit is tuned to the same value. This is possible because the resonant frequency of an LC circuit only depends on the inductance of its inductor and capacitance of its capacitor. Letting $\omega_1 = \omega_2$, the oscillating current through the primary circuit establishes a magnetic field on the interior of the primary’s inductor, which then induces a voltage in the secondary’s inductor. This voltage constitutes a ‘push’ of the secondary circuit. If these ‘pushes’ are well timed—that is, if the time of each push coincides with the peaks of the current in the secondary and is in the same direction—then the secondary coil will resonate. The behavior is just like pushing a child on a swing from both ends: if you push every time the child is at a maximum altitude, you deliver energy to the child’s oscillation, and thus increase the amplitude of his swinging. Similarly, the resonant ‘pushes’ delivered by the primary are synced in such a way that they increase the energy in the oscillations of the secondary. As the energy rises, so does the amplitude of the voltage in the secondary (although it is still oscillating). Eventually, the voltage is high enough that the air around the top load breaks down as charges try to spark to ground (the other plate of the top ‘capacitor’). The spark will ionize the air, turning it to plasma and rapidly increasing its volume, before the volume subsides again. This expansion and contraction of the air’s volume constitutes a pressure wave, which can be interpreted by the human ear as sound, albeit a square, rather than sinusoidal, wave.

Another way to think of a Tesla coil is as a transformer. It takes relatively low voltage AC (15,000V) from the transformer and steps it up to a higher voltage, corresponding to the turn density ratio of the secondary to the primary. Of course, this stepped-up voltage is not enough to cause electrical discharge, which is where resonance comes into effect. The windings are loosely coupled (lots of air space between them), but this is mostly to protect the primary circuit from the induced electric field it experiences from the oscillations in the secondary.

An alternative circuit configuration is shown in Sub-figure 3b. However, in this schematic, high frequency, high voltage oscillations across the high voltage capacitor are mirrored onto the transformer’s secondary
winding. Since we're using a (relatively weak) neon sign transformer, continual operation of a Tesla coil in this configuration will damage the transformer's internal components and is not recommended.

### 4.2.2 Solid State (DRSSTC)

The previous section described the theory of a spark gap Tesla coil. However, a spark gap Tesla coil only operates at one specific frequency. We want to control the frequency of sparks from the top of the secondary. For an A4 note (440Hz), we'd make sparks 440 times per second, and for middle C (262Hz), we'd make sparks 262 times per second.

![Common Tesla Coil Configuration](image1.png)

(a) Common Tesla Coil Configuration

![Alternate Tesla Coil Configuration](image2.png)

(b) Alternate Tesla Coil Configuration

**Figure 3:** Two schematics for a Tesla coil. The top circuit is preferred, as it protects the transformer by shorting across the spark gap. In practice, the secondary coil has many, many more turns (two to three orders of magnitude more) than the primary coil. The toroid acts as a capacitor to ground, in that effectively one plate is the toroid, and the other is the grounding rod or the actual ground.
Therefore we need something that simulates the functionality of the spark gap in a spark gap Tesla coil, but which is controllable by an electrical signal. The best component would be a relay rated for high voltage, high current, and fast switching speeds. However, components like this cost upwards of $2,000, well outside our budget. An alternative approach is to use transistors. The type of transistor best suited to our needs is an IGBT, or Integrated Gate Bipolar Transistor. An IGBT is basically a one-directional switch that closes when the voltage difference between the gate and collector is positive ($V_{GE} > 0$). We’ll likely arrange the IGBTs in an H Bridge pattern to allow current to flow both directions in the LC circuit. IGBT’s are useful because they can withstand both high current and high voltage applications.

4.3 Arduino

What is an Arduino?

There are many types of Arduino boards, such as the Uno, Mega, Leonardo, Due, and more, but all Arduino boards share common features. An Arduino board allows an electrical engineer to interact with a microcontroller (a tiny computer) by uploading programs to be executed and by connecting wires to pins which ultimately lead to the microcontroller. The Arduino board also contains a number of other features, like power indicators, resistors, and more. The essential feature of an Arduino is its ability to execute programs uploaded by a programmer into the microcontroller's memory. These programs are written in the Arduino programming language, which is an extension of C++ that includes various Arduino- and microcontroller-specific features (including Serial, digitalWrite, analogRead, direct register access, and more). Behind the scenes, though, all Arduino code is just C++.

To be technical, the microcontroller doesn’t actually execute the high-level commands that we write in C++ or the ‘Arduino’ programming language; rather, the developer’s computer compiles the ‘Arduino’ code (which is ultimately comprised of C++ constructs) into machine code (pure binary) using avr-gcc, a compiler developed by the Free Software Foundation specifically for use on Atmel microcontrollers (the type found on Arduino boards). avr-gcc is accompanied by avr-libc, a library
of useful C tools, avrdude (AVR Downloader/U-ploaDEr), a project for in-system programming, avr-gdb, a debugger, and more. [17] The compiled machine code is then uploaded onto the board using a USB-to-USB cable. Most boards, including the Arduino Uno and Arduino Mega, use a Standard-A (default computer USB port) to Standard-B converter, but some boards, like the Arduino Due, use a Standard-A to Micro-B converter. An onboard USB-to-Serial converter then converts the incoming USB-encoded data into Serial data that can be fed directly into the microcontroller's ash memory. Additionally, all Arduino boards allow for direct programming of the contents of the microcontroller's memory using the 2x3 pin programming header (known as 'flashing the firmware/software'); however, this technique is dangerous, and should not be used. There are parts of the microcontroller's memory that should not be modified, such as the portion that contains the bootloader, which is a chunk of code that decides when to start executing the uploaded code. (Essentially, it waits for a pause after a stream of incoming data and then runs the uploaded program.) In summary, the Arduino executes a compiled version of an engineer’s C++ code.

For this project I developed software to convert incoming MIDI signals to outgoing electrical pulses (see Section 4.4). This task is challenging. Much like the brain, the Arduino’s job can be separated into three steps: receiving input data, processing data, and sending output data. The theory behind each of the steps is discussed in the following sections.

4.3.1 Serial Input (and Output)

Serial communication refers to the encoding of binary information in a sequence of alternating high voltages and low voltages, corresponding to either a binary 1 (high voltage) or a binary 0 (low voltage). For most Arduino boards, ‘high voltage’ is defined as 5V and corresponds to a binary 1. ‘Low voltage is defined as 0V and corresponds to a binary 0. Some Arduino boards, like the Due, use 3.3V as ‘high’ and 0V as ‘low.’

Every Arduino comes equipped with at least one Serial port, also known as a UART (Universal Asynchronous Receiver/Transmitter) or USART
Serial port is given the name Serial. Serial handles communications in two ways: it can communicate with the computer via USB, and it can communicate with other electrical devices via the digital pins RX (0) and TX (1). RX means receiver and TX means transmitter. [18]

Some Arduino boards come with more than just the standard Serial port. For example, both the Arduino Due and the Arduino Mega 2560 R3 (an improvement over the older Arduino Mega) have additional Serial ports beyond just Serial. These boards have Serial1 on pins 19 (RX) and 18 (TX), Serial2 on pins 17 (RX) and 16 (TX) and Serial3 on pins 15 (RX) and 14 (TX). [19] [20] These extra built-in UARTs are extremely useful because a programmer can employ Serial ports other than RX0 and TX1, which are tied to both Serial and the computer’s USB connection.

It is worth mentioning in passing that a SoftwareSerial library exists that allows arbitrary digital pins to be used as Serial I/O; however, this solution is both slower than the native UARTs built into the Mega’s hardware (as it has to replicate in software an optimized hardware circuit) and has limitations, such as restricted data flow (just one SoftwareSerial can receive data at a given time, and each board has particular limitations as to which pins are available for SoftwareSerial). [21] So, although SoftwareSerial exists, the hardware of the Mega and Due is simpler, faster, and more elegant.

When interfacing with the computer, serial communication is easy. Serial output can be viewed as a sequence of characters via the Arduino environment’s Serial monitor. We use this fact extensively to help debug. Serial input can be passed to the Arduino’s Serial buffer via the text input field at the top of the Serial monitor. While less useful, this feature could be used for debugging in other projects.

Serial I/O also operates on an Arduino’s pins. By default, pin 0 receives data (input), and pin 1 transmits data (output). It’s interesting, although not strictly relevant, that when a serial output call is made (via Serial.println and similar functions), it appears as though the data is only sent
to the computer monitor; however, the serial data is also transmitted on pin 1. This quirk can be used to daisy-chain Arduinos.

An RX pin (such as RX19 for Serial1) reads in data at a set rate (the baud rate) and stores it in a 1024-byte wide Serial buffer. The buffer implements Queue (first in first out) logic. When the RX pin is brought high (at 5V by default) relative to the Arduino’s ground, a logic 1 is appended to a register, and when it is brought low, a 0 is added. When eight bits are added to this register, the entire byte is appended to the Serial buffer and the process resets. Bytes can be drained from the Serial buffer by functions like Serial.read.

Bits in the Serial buffer (and in the Arduino in general) are ordered from least-significant to most-significant. This architecture, shared by most Serial data transfer systems, is referred to as little-endian. Therefore, the sequence of voltages 5V-0V-0V-0V-5V-0V-5V-0V is stored in the buffer as ‘10000110,’ but represents a byte with a value of ‘01100001,’ or 0x61 or 97. This byte can be interpreted as a character using the ASCII encoding (in this case, 97 corresponds to ‘a’), or can be interpreted via another protocol, such as MIDI.

The Arduino’s Serial I/O must have a specified baud rate passed as a parameter to the Serial.begin(int baudRate) command. This value, expressed in units of bps (‘bits per second’), corresponds to the number of times per second the Arduino measures the voltage on an RX pin and adds either a 0 or a 1 to a register. For instance, Serial.begin(1000) will check the RX0 pin every millisecond, but Serial.begin(100000) will check the RX0 pin every 10μs.

4.3.2 Timers

The Arduino usually executes programs synchronously, that is, executing them line by line. However, all Arduinos come with a few timers, which can be used to execute code asynchronously.

A timer is functionally a counter. By default, on every iteration of a clock cycle, a timer’s counter is incremented by 1. When the timer
hits its maximum value, it overflows to 0 and begins the cycle again. Functions can be attached to timers so that when a certain condition is hit, a specified block of code executes. This is a useful tool if one wants to take a measurement exactly every second (for, say, a space launch), or wants to only pulse power to a machine.

Most Arduinos have just 3 timers: Timer0, Timer1, and Timer2; but the Arduino Mega 2560 has six timers: Timer0 through Timer5. This is the key development that will allow us to play five or six notes at once.

A given timer has certain registers associated with it. One of the key aspects of a timer is its maximum size. This size determines the upper bound of what a given timer can count to. Timer0 and Timer2 are 8-bit timers, meaning that they can count to $2^8 = 256$ before overflowing. Timer1, Timer3, Timer4, and Timer5 are all 16-bit timers, meaning that they can count to $2^{16} = 65536$ before overflowing.

Sometimes, the desired period of a timer wouldn’t fit into its counter register (either 1 or 2 bytes) if the counter was incremented on every clock cycle (each clock cycle is $1/62.5\text{ns}$). We certainly couldn’t make a 1Hz timer with increments every 62.5 ns. To solve this problem, there is the notion of a prescaler. A prescaler is a scaling factor that changes the conversion of clock cycles to incrementation. For instance, if Timer1’s prescaler value is 64, the Arduino only increments Timer1’s counter every 64 clock cycles, rather than every one. However, the set of available prescaler values differs for Timer0, Timer2, and Timer1/3/4/5.

We can write our own code to be executed when a timer hits a specified condition; these code blocks are called ISR s, or Interrupt Service Routine s. An ISR is passed an overflow condition, but otherwise works just like a normal function.

Be careful when manipulating timers, as it could cause an accidental malfunction in library procedures. delay() and delayMicroseconds() both rely on Timer0, so if you change Timer0’s prescaler, be sure to divide by the appropriate amount when passing parameters. The Servo
library uses Timer1 extensively, so any changes to Timer1 will affect those calls. Lastly, the tone() function, which generates pulse widths corresponding to given frequencies, uses Timer2.

4.3.3 Port Manipulation

To be a useful component in our assembly, the Arduino must be able to output voltages, quickly turning on and off. The Arduino standard library of functions provides digitalWrite, which is sufficient to set a pin’s output voltage to either HIGH (3.3V on Arduino Due, 5V for all other boards) or LOW (0V) under usual circumstances. However, the conditions we’re operating in are far from usual, and for our purposes, digitalWrite is too slow.

Before we talk more about optimization, we have to talk about clock speed. A processor’s clock speed is a measure of how many instructions the processor can execute per second. For example, my Macbook Pro has a 2.4GHz processor from Intel. The Arduino Mega’s clock speed is 16MHz, meaning it can execute 16 million instructions per second. The Arduino Due’s clock speed is 84MHz, meaning it can execute 84 million instructions per second. But what exactly is an instruction? The definition varies from processor to processor, but think of it as an action the computer performs.

Arduino code, being C++ in disguise, compiles to assembly, and it is this assembly that is uploaded to the microcontrollers. Assembly is a programming language with limited commands that interface directly with the processor. Some commands read from a specific point in memory, while others deal with moving information from memory to registers and back. Assembly can be directly translated to binary code (machine code). Each assembly command takes a specific number of clock cycles to execute. For instance, the Arduino assembly commands cbi (clear bit i) and sbi (set bit i) both take two clock cycles. Jumping back to the start of a while loop is an rjmp (relative jump) and also takes two clock cycles. So, if we want to maximize the switching speed of the Arduino we have to minimize the number of clock cycles a single switch takes to execute.
digitalWrite takes two parameters: an integer corresponding to the pin of interest, and either HIGH or LOW specifying whether to bring the pin to 5V or 0V. However, the internal code of digitalWrite() has a slow chain of conditional checking to extract the internal pin number from the pin parameter. Therefore, there must be a faster way.

There is! Each pin of the Arduino is associated with a particular bit in three different hardware registers: PORTx, PINx, and DDRx, where x is a letter between A and L (on the ATmega2560—older micro controllers have fewer registers). Each register is eight bits (1 byte) long, and therefore each register contains information about eight different pins. For instance, information about Pin D5 is contained in the fifth bit from the right (fifth least significant) in registers PORTD, PIND, and DDRD.

DDRx contains information about pin configurations. A 1 means that the pin is configured as input; a 0 means it is configured as output. PORTx controls whether a pin is HIGH or LOW, and PINx reads the status of pins that were configured as inputs using DDRx. Both DDRx and PORTx are read/write, but PINx is usually read only. However, the ATmega2560 allows writing a logic 1 bit to a PINx register to toggle the state of the corresponding bit in the corresponding PORTx. That is exactly what we need in order to toggle pulses to the primary circuit.

There are benefits and drawbacks to using direct port manipulation. The code is harder to debug and much harder for an outside programmer to understand. It is much less portable than digitalWrite and digitalRead, since the pin mappings are different on the ATmega2560, the ATmega328, the ATmega168, and the ATmega8. Additionally, it is very easy to cause malfunctions in built-in systems, especially when mistakenly setting registers directly, rather than and-ing or or-ing them with given bitstrings.

The biggest benefit for our purposes is the fast speed of direct port access compared to the clunky digitalWrite. Additionally, direct port access allows multiple pins to be set at exactly the same time (not
useful to us, but interesting nonetheless). Finally, this approach can reduce program size in the event that the program is almost too large to fit in the board's memory.

The ATmega2560, the processor which underlies the Arduino Mega 2560, has the pin mappings specified in Figure 4. Some pins have an internal name of Pxn, where x is a letter between A and L, and n is a number between 0 and 7. These pins correspond to the nth bit from the right in registers PORTx, PINx, and DDRx.

We've discussed the pin mappings for the ATmega2560 because they are easier to understand and because they are more analogous to the popular Arduino Uno. However, the Arduino Due actually uses an entirely different setup to toggle output states, which we will discuss briefly. The Arduino Due provides a plethora of 4-byte registers (it's main advantage over the older Arduino models) that control the input and output of various pins.

The Arduino Due has analogous registers to DDRx and PORTx. We'll ignore PINx for now. The REG_PIOx_OER, or Register Parallel Input/Output X Output Enable Register, determines whether a pin is configured as input or output. REG PIOx SODR, or Set Output Data Register, will set a pin high if the corresponding data bit is brought high. REG PIOx CODR, or Clear Output Data Register, will set a pin low if the corresponding data bit is brought high. The abstraction of PORTx and PINx into x SODR and x CODR allows a more sensible treatment of pin output values. Instead of using PORTx and PINx as direct access to the pin's output values, the programmer can write instructions in the SODR and CODR register which then carry over to the actual output pins. ‘x’ can be any letter between A and F.
Figure 4: Pin mapping for the ATmega2560, the controller that underlies the Arduino Mega 2560.
### 4.4 MIDI

#### 4.4.1 What is it?

MIDI, or Musical Instrument Digital Interface, is a standard that allows two electronic instruments to communicate. It was developed around the time electronic music started surfacing in the 1970s. Various manufacturers formed the MIDI Manufacturers Association (MMA), a coalition that has created two standards—General MIDI Level 1 (GM1), and General MIDI Level 2 (GM2)—while still allowing companies to develop their own unique version through the use of System Exclusive messages.

GM1 specifies the use of MIDI events to communicate information in the form of instructions. It specifies 128 instruments (denoted as ‘programs’) and 64 percussion sound patches. It also specifies a set of controller actions.

MIDI has 16 channels of communication on which it can send or receive data. Usually, however, Channel 10 is set aside for percussion tracks. [24]

#### 4.4.2 Information Protocol

The basic unit of MIDI communication is a MIDI event, which is essentially an instruction. This event is normally a set of three bytes which encodes four pieces of information: the command byte, the first data byte, and the second data byte. The command byte is broken into four bits, which represent a status value and a channel value. The status value communicates which type of MIDI message follows, and

<table>
<thead>
<tr>
<th>Name</th>
<th>Range</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Status</td>
<td>0x0-0xF</td>
<td>Type of MIDI message</td>
</tr>
<tr>
<td>Channel</td>
<td>0x0-0xF</td>
<td>Channel number</td>
</tr>
<tr>
<td>Data Byte 1</td>
<td>0x0-0xFF</td>
<td>First data byte (see Table 2)</td>
</tr>
<tr>
<td>Data Byte 2</td>
<td>0x0-0xFF</td>
<td>Second data byte (see Table 2)</td>
</tr>
</tbody>
</table>

Table 1: A description of the bytes found in a typical MIDI event transmission
the channel value communicates which channel the event should influence. This information is summarized in Table 1.

<table>
<thead>
<tr>
<th>Status</th>
<th>Value 1</th>
<th>Value 2</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>Note</td>
<td>Velocity</td>
<td>Note o</td>
</tr>
<tr>
<td>9</td>
<td>Note</td>
<td>Velocity</td>
<td>Note on (if velocity = 0, note o )</td>
</tr>
<tr>
<td>A</td>
<td>Note</td>
<td>Value</td>
<td>Polyphonic pressure</td>
</tr>
<tr>
<td>B</td>
<td>Controller #</td>
<td>Value</td>
<td>Controller change</td>
</tr>
<tr>
<td>C</td>
<td>Program</td>
<td>-</td>
<td>Program change</td>
</tr>
<tr>
<td>D</td>
<td>Value</td>
<td>-</td>
<td>Channel pressure</td>
</tr>
<tr>
<td>E</td>
<td>Value</td>
<td>Optional Value</td>
<td>Pitch bend</td>
</tr>
<tr>
<td>F</td>
<td>Value</td>
<td>Optional Value</td>
<td>System Exclusive (SysEx) or Realtime Message</td>
</tr>
</tbody>
</table>

Table 2: A description of the data bytes expected by the possible status values.

The first bit of any incoming MIDI message is a 1, so the range of possible status values is really 8 to F. The status value determines what type of MIDI message the incoming event is, and also determines how the data bytes should be interpreted. This information is shown in Table 2. The channel values range from 0 to F, specifying any of the 16 possible channels of MIDI communications (ranging from '0000,' which is channel 1, to '1111,' which is channel 16). Therefore, the entire command byte ranges from 0x80 to 0xFF (128 - 255).

For pitch bend (E), if a machine sends two value bytes, they are interpreted as a LSB and MSB, whereas if it sends just one, that byte is interpreted as a MSB.

Note that some status commands (C and D, and sometimes E and F) are associated with fewer than one data byte. However, a majority of the events that are sent are either note on or note off events, both of which require two data bytes. We’ll focus on these events at length.

For both note on and note off events, the first data byte corresponds to the note pitch, and the second data byte corresponds to the velocity, or
volume, of the note. A note off command always has velocity zero, but some electronic instruments transmit note on messages with a velocity of zero to indicate that the note should be turned off.

The MIDI note value ranges from 0 to 7F (0 to 127), as does the velocity, where a higher value corresponds to a louder volume. The MIDI pitch value is associated with musical notes via Table 3. [24]

4.4.3 Decoding MIDI

The Arduino has to convert the sequences of bits it gets from a MIDI stream to sensible information about the MIDI event. To further understand how to convert Serial data to MIDI events, we’ll practice with a few sequences of bits.

First, let’s pretend the MIDI stream sends the following data bits in sequence:

100100010011010001101011

The first step is to convert these to bytes: our command byte (1001 0001) is 0x91, or 145; our first data byte (00110100) is 0x34, or 52; and our second data byte (01101011) is 0x6B, or 107. The 9 (status value) in the command byte indicates that this MIDI event is a note on event. The 1 (channel value) in the command byte indicates that this MIDI event takes place on channel 2. Since this is a note on command, the first data byte corresponds to a MIDI note pitch of 52, which (using Table 3) means E4. Finally, the second data byte corresponds to a velocity of 107, which is about 84% of the maximum. Putting it all together, we conclude that this MIDI event is an instruction to play E4 on channel 2 at 84% volume.
Table 3: Translation table between MIDI note values (as received by a note on or note off message) and musical notes.

Let’s now dive into the hardware, looking at real voltages. MIDI operates at a baud rate of 31,250 bps. This means that each bit has a duration of exactly 32μs. MIDI uses the 8-N-1 data encoding protocol, which specifies exactly how data bits are transferred over a wire. (Note that the phrase ‘data bit’ is being used in a more general context here than above.) In the 8-N-1 protocol, each byte of information is associated with more than 8 voltage ‘bits.’ For every eight data bits (one byte), there are no parity bits (the ‘N’ in 8-N-1) and 1 stop bit. All Serial communications are prefaced with a start bit to inform the listener (in our case, the Arduino) that new data is arriving. Therefore, each MIDI byte is really encoded in ten bits: a start bit (low), eight data bits, and a stop bit (high). Therefore A MIDI source can transmit at most 3,125 bytes every second.

MIDI has no error-detection or error-correction capabilities due to a lack of parity bits, and therefore the MMA requires that the maximum cable length be 50 feet to minimize the degradation of the signal due to random noise pickup.

One last note about the MIDI protocol. Bits are transmitted (as in almost all Serial systems) from least significant bit to most significant
bit. This may appear ‘backwards,’ since we often like to write bytes from most significant bit to least significant bit. This quirk is a feature that allows computers to process data more quickly.

Figure 5 shows a real oscilloscope graph of voltage on an Arduino’s RX pin relative to the Arduino’s ground. Let’s try to determine what MIDI event was sent. Each box is 100\(\mu\text{s}\) wide, so each bit is a little less than two ticks wide.

The problem is left as an exercise for the reader.

**Figure 5:** Voltage incident on the Arduino’s RX pin when a specific MIDI event is sent. Try to figure out what event was sent.

Software has been developed to decode incoming MIDI signals in the exact same way we just did.

Conventional MIDI cables consist of five pins in a half-circle configuration. If looking directly at a male MIDI plug, these pins are, from left to right: 1, 4, 2, 5, 3. Pins 1 and 3 are not used, and only exist for forward-compatibility in MIDI hardware.
MIDI hardware is not capable of both transmitting and receiving messages on the same cable, so often MIDI contains two cables, one to transmit and one to receive. For the transmitter, pin 4 is always at 5V, and pin 5 is either at 0V or 5V. On the receiver the same is true, with the addition that pin 2 is grounded. This ground provides the shielding for the signal pair of wires.

4.5 Circuit Elements

In addition to the conventional circuit elements described previously, we use unconventional components in our circuitry. The following sections describe the theory of each of these components. Specifications are described later on, when we discuss the practical application of these idealized components to our overall goal of creating a musical Tesla coil.

A note about active and passive components: passive components are those that do not require an external voltage supply (more than just the circuit connections) to function. Examples of passive components include resistors, capacitors, diodes, switches, etc. Active components require an external power source, conventionally denoted $V_{cc}$, originally named after the common collector of an NPN transistor based circuit. This voltage supply usually powers internally logic components and other active components inside the overall active component.

4.5.1 Opto-isolator

An opto-isolator is a circuit element that connects two electrically disconnected circuits via optical transmission. A photodiode (think LED) on one side of the circuit will transmit a signal if and only if there is a voltage present across the ends. This light is fiber-optically transferred to a photo-transistor on the other circuit that activates when the light is detected. Many opto-isolators employ internal mechanics different from these, but this is the general idea. A sample schematic is shown in Figure 6.
Opto-isolators are extremely useful because they allow two circuits to have separate grounds. We’ll use this fact to isolate the two halves of our MIDI circuit.

4.5.2 Fiber Optical Transmitter and Receiver

Electricity is not the only way to send signals. We can also send signals via light. To do so, we need a device that can convert electrical signals to light signals and a device that can convert light signals back to electrical signals.

An LED (Light Emitting Diode) is one such component. Given an applied voltage across its leads (past some threshold voltage), an LED will light up. If the voltage difference is too high, the LED will burn out as the physical and chemical structures responsible for the emission of light begin to break down. Since an LED is a diode, it only allows current through in one direction, so the polarity of LEDs matters. All fiber optic transmitters are essentially glorified LEDs—perhaps they can turn on and off very quickly, or perhaps their spectral output is less broadband—but fundamentally a fiber optic transmitter can be thought of as a fancy LED.

On the other hand, we’ll need some device to sense an incident light source and produce some electrical signal. These components are called phototransistors. All semiconductor materials respond to a wide range of frequencies, some more than others. When light of any wavelength hits a material with excess electrons (N-type) or electron holes (P-type), it alters the concentration of charge carries in the material, thus changing the material’s electrical properties, specifically with regards to conduction. All semiconductor materials respond to applied light in this way, but the largest response is to infrared light (about 800nm, in the near-infrared portion of the electromagnetic spectrum). In fact, any transistor can be converted into a phototransistor with a transparent case and an extremely strong source of IR. Phototransistors on the market use focusing lenses, doping, and other techniques to amplify this natural effect to the point where the final component is effectively
a light-controlled on/off switch, although in reality it demonstrates the
same resistive heating that normal transistors do when switching states.
[25] As with the fiber optic transmitter, these fiber optic receivers are
often tuned for desired parameters, such as switching speed, minimum
incident light, or maximum output load.

![Figure 6: Internal Circuit Diagram for the CNZ3731 opto-isolator. We are not
using this opto-isolator (although we do have it in the lab) but its schematic is
relatively straightforward.](image)

We still need some way to transmit light information from our fiber optic
transmitter (a specialized LED) to the fiber optic receiver (a phototransistor).
A fiber optic cable is a specialized flexible cable that traps light inside for
very long distances. Google Fiber is using similar technology to transmit
data at up to 1000 Mbps, or 125 megabytes per second. [26]

### 4.5.3 Voltage Regulator

A voltage regulator is an electrical component that accepts some variable
voltage and outputs a specific voltage. For example, a voltage regulator
could accept a fluctuating voltage—perhaps 19V plus or minus 0.5V—
and output a consistent 15V. Voltage regulators rely on a feedback
mechanism to keep modifying the output voltage until it exactly matches
a reference voltage stored in the internals of the regulator.

While some voltage regulators use electromechanical feedback to
control the output voltage, our voltage regulators use a feedback loop
consisting of an op-amp connected to a Darlington pair of transistors
(which enhances current gain) to modify the output. An example circuit of a voltage regulator is shown in Figure 7.

4.5.4 NOT Gate

A NOT gate performs a unary logical not operation on an input signal. Essentially, if the input is low, the output is high, and if the output is high, then the input is low. Each NOT gate defines its own cutoff between a low and a high input signal. A simple NOT gate needs four pins: \( V_{cc} \), GND, input, and output. Specifically, if the input voltage is below some threshold, the output will be at (approximately) \( V_{cc} \), and if the input voltage is above some threshold, the output voltage will be at GND. A NOT gate can be constructed with one transistor. If there is sufficient voltage at the base of our bipolar junction transistor, then the transistor will act as a closed switch, allowing voltage to drain to ground. However, if there is no voltage at the base of the transistor, then the transistor will act as an open switch and the output will be at a nonzero voltage.

The class of circuits that uses BJTs and resistors in this way is known as TTL, or transistor-transistor logic. The other large class of circuits is CMOS, or complementary metal-oxide-semiconductor, which uses a different configuration of semiconductor devices to perform logic and amplification. While other classes exist (RTL, resistor-transistor logic, and DTL, diode-transistor logic, to name a few), TTL and CMOS are by far the most prevalent. Even though there are fundamental differences between TTL and CMOS circuits, we’ll treat them as identical with regards to the overall logic performed on input signals. [27] [28] [29] Technically, all our ICs use TTL.
More complicated logic gates exist, performing operations such as NAND, OR, AND, and many more. Binary logic operations such as these will require two transistors to handle two inputs. These logic gates can be combined in various ways to produce adders, memory, and many other high-level components relevant to a modern computer. In fact, it is possible (although hardly advisable) to make a computer entirely out of NAND gates.

4.5.5 D-Type Flip-Flop

A flip-flop is an active, stateful electrical component—that is, its output at any point in time is determined not only by its inputs at that moment, but also by previous inputs. In this way, the flip-flop is a piece of memory. Specifically, the flip-flop can be in one of two states: high or low. It is possible to build a memory unit, perhaps for a computer, using flip-flops, since each one can store and retain one bit of information. Although the flip-flop is internally just a complicated circuit built of transistors, resistors, and other electrical components, we will treat it as a black box which has memory. It is conventional for flip-flops to provide both an output pin and its complement, to reduce the unnecessary use of NOT gates.

Conventional (SR) flip-flops have two inputs: a set (S) and a reset (R). If the R pin is brought high, the flip-flop is brought low. If the S pin
is brought high, the flip-flop is brought high too. If both pins are low, then the flip-flop maintains its state. Internally, SR flip-flops are cross-coupled NOR gates. Draw the circuit diagram to verify that the extra inputs to the NOR gates for a set-reset pair.

We use a D-type flip-flop, which is slightly more complex. In addition to the set and reset lines, there is both a clock line and a data line (hence the D). At a specific point in the clock's cycle (perhaps on the rising edge), the value of the data line is mirrored to the state of the flip-flop, which can be read as the flip-flop's output. The set and reset lines on a D-type flip-flop work the same way they do on an SR flip-flop, and are often used to force the flip-flop into a particular state, regardless of the clock or data lines.

4.5.6 Integrated Gate Bipolar Transistor

An integrated gate bipolar transistor (abbreviated IGBT) is a transistor designed for switching quickly at high voltages and currents. It uses a particular arrangement of semiconductor materials to achieve this goal. Notably, instead of the three regions of the familiar bipolar junction transistor (PNP or NPN), IGBTs have four alternating regions (NPNP or PNPN). It is this configuration that gives IGBTs their switching power.

4.5.7 Gate Drive Transformer

The gate drive transformer we use is a 1:1:1 voltage transformer that allows a voltage difference across its input leads to be mirrored across two pairs of output leads. A 1:1:1 gate drive transformer consists of three wires that are twisted around each other, then spiraled through a metal toroid. Just as in any other transformer, the changing current through the input wire creates a changing magnetic field in the metal toroid. This changing magnetic field induces an electromotive force, which essentially acts as a voltage, in the pair of output wires. Since the coil ratio is 1:1:1, the voltage ratio is also 1:1:1.
4.5.8 Current Transformer

A current transformer decreases the magnitude of alternating current in a circuit. Current transformers are often used to measure the frequency of high voltage alternating current, which cannot be fed directly into measuring instruments. Usually the lowered current is used directly for measurement; however, we will use it as a feedback transformer so that the logic circuit can 'sense' the oscillating current in the power circuit. A current transformer works using Ampère's Law and Faraday's Law. As current alternates in the measured wire, a magnetic field is established in the current transformer's core. This alternating magnetic field induces an electromotive force in the measuring wire, which then drives a proportional current (although phase shifted by a quarter of a cycle).

5 Design

The design of this project is separable into three distinct sections. The first is the experimental design of the project itself—its construction as the union of independent modules. The second is the design of the MIDI to Arduino circuit. The last is the design of the primary Tesla coil circuit. Some physical elements were designed and implemented, but to such a small degree as to be uninteresting.

Overview

The project is designed to break apart into manageable, independent modules. These modules are shown in Table 4.
From | To | Means
---|---|---
Piano/Computer | MIDI Stream | Built in to Q61 Keyboard
MIDI Stream | Arduino Serial In | Optocoupler circuit
Arduino Serial In | MIDI events | Code (MIDI library)
MIDI Events | Timers | Code
Timer Handlers | Electrical Pulses | Code (direct port access)
Electrical Pulses | Primary Circuit (logic half) | Transistors/Fiber Optics
Primary Circuit (logic half) | Primary Circuit (power half) | Many logic and power ICs
Primary Circuit (power half) | Resonance in Secondary | Tuned LC Circuits
Voltage in Secondary | Electrical Arcs | Breakdown

| Table 4: A breakdown of the units necessary for a successful project. |

With this design, each small problem could be tackled individually, then merged together to form a successful final project. A rough flow chart of the project is shown in Figure 8.

5.1 MIDI to Arduino Circuit

The goal of the MIDI to Arduino circuit is to convert incoming MIDI signals into Serial data that the Arduino can read and understand. In this project, we only use the MIDI output of an Alesis Q61 piano keyboard or a computer, but in theory the incoming MIDI data stream could come from any MIDI-enabled instrument. In order to use a computer as a MIDI source for the Arduino, a USB-to-MIDI cable must be employed, since no computers have native MIDI out ports.

Incoming MIDI data, transmitted over a MIDI-to-MIDI or a USB-to-MIDI cable, is converted into a breadboard-compatible format by a female MIDI connector (5 pin DIN 180° connector) from SparkFun. This component allows male MIDI cables to be directly ‘plugged into’ a breadboard without the need for extraneous soldering. [30] The MIDI connector we use is shown in Figure 9. As per the datasheet, the pins correspond to MIDI connections 3, 5, 2, 4, 1 from left to right when viewed from the front with the pins in a convex up (smiling) configuration. [31]
Figure 8: Abstract diagram of the flow of data in the project and the tools used to process it.

(a) Front view

(b) Rear view

Figure 9: Two angles of the SparkFun breadboard-friendly MIDI connector.
The rest of the circuit roughly follows the MIDI to DIN electrical specifications, published by the MIDI Manufacturers Association. Their circuit is shown in Figure 10.

Our slightly modified circuit is shown in Figure 11. We have only changed the diode (from 1N914 to 1N4148), the optocoupler (from PC-900 to 6N138) and the value of the drain resistor (280Ω to 470Ω).

We use an optocoupler to isolate the MIDI side of the circuit from the Arduino side of the circuit for multiple reasons. First, the optocoupler allows us to maintain separate grounds on the MIDI and Arduino side of the circuit. Since a MIDI signal consists of a series of voltage spikes, if the signal and receiver circuits are not optically isolated, loops of current through the common ground will corrupt the signal.

Another benefit of the optocoupler is to x any signal degradation that may have occurred over the MIDI cable. After all, MIDI provides no error-correction procedure in its protocol, so random fluctuations may (and do!) significantly interfere with the quality of the signal. Therefore the optocoupler not only removes minor noise fluctuations but also brings the 'high' voltage of the MIDI signal to a constant amount (see Section 6.1).

After trying a variety of optocouplers, we settled on the 6N138 optocoupler. Although it is an active component, its switching speed is extremely fast (less than 10μs for both rise and fall). On the master side, pin 2 is the cathode of the internal photodiode (LED), and pin 3 is the anode. On the slave side, pin 8 is power, pin 7 can increase switching speed if it has a pull-up resistor, pin 6 is the output, and pin 5 is ground. The arrangement of the Darlington pair of transistors in this component ensures that pin 6 will be at ground if and only if there is a voltage across pins 2 and 3. The schematic for the 6N138 is shown in Figure 12.
Figure 10: Electrical specification from the MIDI Manufacturers Association for the conversion of a MIDI In signal to a UART (top circuit). Specifications are also given for MIDI Thru (middle circuit) and MIDI Out (bottom circuit).

Figure 11: Schematic of the MIDI to Arduino circuit. The MIDI transmission is carried on pins 4 and 5.
In Figure 11, R1 drops the voltage across the optocoupler's inputs in order to protect the photodiode (LED) in the rare case that the MIDI provider does not adhere to the suggested standard of providing MIDI power from 5V through a 220 resistor. The maximum recommended voltage across the LED is 1.7V, and it can break down at as little as -5V, so the resistor keeps the LED from breaking. [33] The photodiodes of other optocouplers may break down at even lower voltages, so the resistor is a good safety precaution, although not strictly speaking necessary. The 1N4148 diode is a fast-switching diode capable of recovering from forward conduction in at most 4 nanoseconds. It has replaced the 1N914 originally suggested by the MIDI Manufacturers Association, because it has the same electrical specifications but slightly lower leakage current (25nA vs. 5μA) and a slightly higher maximum average forward current (150mA vs 75mA). [34] We use this diode in order to minimize feedback loops which distort the MIDI signal, and to protect the photodiode of the optocoupler from improperly wired MIDI pins, where pin 5 is always at a higher voltage than pin 4.

The optocoupler setup maintains the logic of MIDI pin 5, which carries the MIDI data. If the MIDI source sends a logic 1, MIDI pin 5 will be at 5V, so there will be no voltage difference between pin 2 and pin 3 of the optocoupler. Therefore, pin 6 of the optocoupler will not be connected to ground, and the Arduino's RX pin will read a high voltage (by way of R2), corresponding to the MIDI source's original logic 1. We increased the value of the drain resistor because the Arduino Due only accepts voltages on I/O pins between 0 and 3.3V. By increasing the value of the drain resistor, we lower the voltage incident on the Arduino's RX pin (which is an I/O pin) so that we are less likely to damage the Arduino's microcontroller. Conversely, if the MIDI source sends a logic 0, MIDI pin 5 will be at 0V, so there will be a voltage difference across the photodiode. Then, pin 6 of the optocoupler will be connected to ground (through the ‘inner’ transistor in the Darlington pair), and so the Arduino’s RX pin will see 0V, correctly corresponding to the original logic 0 from the MIDI source.

The incoming Serial data is processed by a MIDI library, which calls function handlers in response to different MIDI status values.
Figure 12: Schematic for the 6N138 Optocoupler. Pin 2 is the cathode, and pin 3 is the anode of the control circuit. On the slave circuit side, pin 8 is power, pin 5 is ground, pin 6 is output voltage, and pin 7 (optional connection) is a pull-up that increases switching speed. Pin 6 can drain to ground if and only if there is a voltage across the input LED.

5.2 Primary Circuit

The overall goal of the primary circuit is to create a series of electrical breakdowns (from the top conductor to ground) exactly when a signal is on. To repeat, we want to be able to control exactly when arcs occur by sending electrical signals.
The original design for this Tesla coil’s primary circuit is an adaptation of the circuit used by one Tesla (see Figure 13).

The overall circuit is comprised of two halves. There is a power half, which contains high voltage and high frequency electricity, and a logic half, which controls the current in the power half. We’ll begin by looking at the power half, and conclude by discussing the logic half.

5.2.1 Power Circuit

The primary circuit’s power half consists of two parts: (1) a voltage double and rectifier, and (2) a half bridge.

Ultimately, the Tesla coil is powered by electricity from a wall socket. There are usually three holes in an electrical socket. These three pins are associated with hot, neutral, and ground. The hot connection fluctuates at 120V AC at 60HZ. The neutral connection is held at zero voltage, referenced to Earth’s voltage. The ground pin is optional and just for safety; it is frequently connected directly to ground via a stake in the ground or something similar. The neutral pin is used to carry current back to ground, whereas the ground pin is not.

Let’s take a closer look at the voltage double and rectifier. When the AC power supply voltage is negative, the top diode conducts and the top capacitor charges. When the AC power supply voltage is positive, the bottom diode conducts and the bottom capacitor charges. The voltage drop across each capacitor is equal to the voltage of the AC power supply, so the overall voltage drop across both capacitors (which are in series) is twice the source voltage—hence, a doubler. Note that the output voltage measured across the resistive load will still pulsate, although all the voltage is in the same direction and the maximum voltage is twice the source voltage of our power supply.

We use 2 MUR460 power diodes (rated at 60A, 1000V DC) to rectify the AC, and two large capacitors to smooth out the bumps (3188FH222M350AMA2 - 2200 uF +- 20%, 250VDC, 400 surge, 85 max c ambient).
The half bridge is what directly drives the oscillations in the primary circuit that allow the Tesla coil to build up. In theory, we pulse the DC voltage across the LC circuit, initiating oscillations. We alternate the direction of the voltage across the tank capacitor and inductor to induce AC current in the primary circuit, much like the oscillations in an LC circuit. For this to work, we need to switch the DC voltage at the resonant frequency of the Tesla coil.

Figure 13: Schematic published by oneTesla for the primary circuit. Before the current transformer is a rectifier and doubler circuit that converts 120V AC to 240V DC.

5.2.2 Logic Circuit

The logic circuit is divided into two parts. There is a signal side and an output side.

The two inputs to the logic circuit are an interrupter signal and a feedback signal. The interrupter signal is from the output of the fiber-optic receiver, and corresponds to the electrical pulses sent out by the Arduino. The feedback signal is from a current transformer jumped off of the main primary circuit (the middle line in the half bridge). A combination of the interrupter signal and feedback signal controls the output of a D-type flip-flop, which then determines the output
states of two complementary gate drivers. The differential voltage signal between the two gate drivers is routed through a 1:1:1 gate drive transformer, which allows easy switching of the transistors.

The feedback signal is driven by the output of a 300:1 Triad current transformer, which allows us to ‘keep an eye on’ the oscillating current in the primary circuit.

The AC signal is squared up by sending it twice through an inverter, and this signal is fed into our flip-flop in conjunction with the inverted interrupter signal. The clocked AC signal is also sent to the inputs of two complementary gate drivers, and the enable line of the drivers is fed by the output of the flip-flop. The differential voltage from the output of the drivers powers the gate drive transformer.

5.3 Physical Design

Many of the existing physical objects are left over from a previous ASR project. Batchelder’s components emphasized interchangeability, which makes modifying his system easier. While I left most of his components intact, I modified the existing primary coil. It was very crooked, with sloppy soldering that allowed sparks to short-circuit to the secondary (through insulation!) from the primary. To that end, I am prototyping a 3D-model of a holder for the primary coil, which locks in place (preferably locks into the table) and adds structural integrity to the primary coil. Since the 3D printer’s platform is too small for the entire piece to be printed at once, I’m creating separate interlocking parts. The prototype is shown in Figure 14. Since I am still determining measurements, this diagram has no specific dimension and is only intended to give an idea of what the supports will look like.
Also, time permitting, I’d like to redesign the Tesla coil stand to make it more stable. As of now, one of the table’s legs is not physically attached to the tabletop and many connections are made by epoxy rather than screws.

6 Results

6.1 MIDI to Arduino

6.1.1 MIDI Before and After Optocoupler

The optocoupler performed admirably in cleaning up the incoming MIDI signal by reducing noise, raising the reference voltage while still maintaining the different bits due to its very fast switching speed.
For consistency in this section, all graphed bright white lines represent the voltage difference between MIDI pins 4 and 5 at the MIDI connector (before the optocoupler), and all graphed gray lines represent the voltage difference between the output (pin 6) of the optocoupler and the Arduino’s ground (after the optocoupler).

Figure 15 shows the impact of the optocoupler on the signal quality. Miscellaneous noise is removed, and the max voltage is raised to a level the Arduino can recognize as a binary 1. Note the inverting action of the optocoupler circuit: when the voltage difference between MIDI pins 4 and 5 is low, the optocoupler outputs a high voltage, as expected.

Figure 16 encapsulates the fast logic of the optocoupler. From Sub-figure 16a, we see that the propagation delay from the start of a rising signal edge to the start of a falling output edge is 2.6μs. Sub-figure 16b shows that the propagation delay from the start of a falling signal edge to the start of a rising output edge is 1.8μs.

Figure 17 shows the falling and rising durations for the output of the optocoupler. Sub-figure 17a and sub-figure 17b give the falling and rising times as 300ns and 3μs respectively.

Combining our results, we see that the overall propagation time to a logical low (\(T_{PHL}\)) is 2.6+0.3 = 2.9μs, and the overall propagation time to a logical high (\(T_{PLH}\)) is 1.8+3 = 4.8μs. These values are consistent with the typical values of TP HL = 1.5μs and TP LH = 7μs for the 6N138 optocoupler, specified on page 4 of the optocoupler’s data sheet. [33] One reason our values differ from the specifications is that, as described in Figure 24 of the datasheet, the cuto for both ‘sufficiently low’ and ‘sufficiently high’ is 1.5V, whereas our tests consider ‘sufficiently low’ to be 0V and ‘sufficiently high’ to be 4V. If we use Fairchild’s cutoffs to estimate our optocoupler’s switching speeds, we find \(T_{PHL} = 2.5\mu s\) and \(T_{PLH} = 2.4\mu s\), which is more self-consistent. Either way, the optocoupler is clearly switching incredibly quickly. It is this high-speed switching that allows for the creation of such high-quality MIDI data for the Arduino to read.
6.1.2 Arduino Reads MIDI Data

The software works exactly as intended. Using the MIDI library, the Arduino was able to read in a sequence of voltages and call a desired function, such as HandleNoteOn or HandleSystemReset according to the type of MIDI signal. Sample Arduino Serial output might be NOTE ON: Channel 1, Pitch 47, Velocity 78. Computed Frequency 123.47Hz.

![Image of incoming MIDI data before and after the optocoupler.](image1)

**Figure 15:** Incoming MIDI data before and after the optocoupler.

(a) Rising Signal Edge Propagation Delay
(b) Falling Signal Edge Propagation Delay

![Image of signal propagation delay of optocoupler.](image2)

**Figure 16:** Signal propagation delay of optocoupler. The voltage and time scales are identical.
6.2 Arduino In (MIDI) to Arduino Out (PWM)

All of the optimizations we made in the Arduino code were effective. These results are the most satisfactory, because they prove that all of the code came together to function as desired.

6.2.1 Processing Delay

The biggest speed challenge was in interpreting a sequence of high and low voltages as a MIDI message and starting a timer at the appropriate frequency. Our code utilizes the power of the MIDI library as well as the Due Timer library to quickly process data and start the timer. Figure 18 shows the overall delay between the incoming MIDI message and the first outgoing electrical pulse. Shockingly, the Arduino was able to respond to the incident stream of voltages in just 112μs! However, this defines the upper bound of the possible delay, because the final MIDI data byte in reality ends with a high (1) stop bit. Since the MIDI protocol operates at 31,250 bps, this final high bit (which is ‘hidden’ in the post-MIDI voltage stream) has a duration of 32μs. At what point during the 32 s-long high bit does the Arduino read in the voltage as a digital 1 into the Serial input buffer? We have no idea. As far as our research suggests, there is no documentation on precisely when the Arduino reads in the final high stop bit. Therefore, the processing delay could be as low as 80μs! If we assume the Arduino reads in the voltage
value uniformly at random from the \(32\,\mu s\) interval, then the expected value of the processing delay is \(96\,\mu s\), still less than \(100\,\mu s\).

This processing delay, however, takes into account the delay incurred both by the MIDI library and by the Due Timer library. Figure 19 shows the processing delay to interpret the MIDI signal and call the associated function handler (tasks performed by the MIDI library). The blinkLED method was called from within the body of the NoteOn method, albeit after three if statements. Therefore, the visible delay represents the time taken to completely analyze and interpret the incoming MIDI signal. As before, we don’t know when in the last \(32\,\mu s\) the Arduino actually finishes reading the final MIDI data byte (returning it to a Serial.read call)—hence the lack of cursors; however, we can approximate that the MIDI processing code took between \(0\,\mu s\) and about \(20\,\mu s\) to execute. Again, if we assume the Arduino read in the code at the halfway point, we get an expected processing duration of \(4\,\mu s\)—extremely fast!

![Figure 18](image.png)

Figure 18: The processing delay of the Arduino from reading in a MIDI signal to the initiation of a timer’s first generated output pulse. Since the Arduino could read the final high stop bit anywhere in the last bit’s time interval, the delay could be as little as \(80\,\mu s\).
Note: The following tests were conducted with the user output disabled (useSerial=False; useLCD=False;). If the booleans useSerial or useLCD are set to True, the processing delay will be much greater (empirically about 10 to 15ms, depending on the length of the output message). Since this latency is incurred only once (when starting the note), it doesn’t matter much to the overall sound. Nevertheless, it is satisfying to see the rapid response of the code.

Figure 19: The processing delay between a MIDI signal and the execution of an associated function.
6.2.2 Characteristics of Output Pulse

We programmed the Arduino to generate a sequence of output pulses that are (a) a specific duration and (b) as square as possible. For (a), we used the Arduino’s built-in delayMicroseconds function, which is accurate to values as low as 3 μs, at least in Atmel-based microcontrollers. Figure 20 gives an overview of the pulse generated by the Arduino as well as a measurement of its width: exactly 10μs, as expected.

The most impressive feature of these generated pulses is the sharp rise and fall edges. Figure 21 shows a zoomed-in version, giving an approximate measurement of the rise and fall times on the edges of this pulse. The Arduino takes 150ns to switch from a stable low output voltage to a stable high output voltage or vice versa. However, if we look even more closely, as in Figure 22, we see that the initial rise time (before stabilization) is 12ns, and the initial fall time (before stabilization) is 14ns. At the Arduino Due’s clock speed of 84MHz, these blisteringly
fast switches correspond to roughly one clock cycle ($\frac{1}{84,000,000} = 11.9$ns). Therefore, these figures are proof that direct port manipulation is an extremely fast alternative to the clunky digitalWrite, which can take up to $20\mu$s to execute.

![Image of waveforms](image.png)

(a) Stable rise time of generated pulse. (b) Stable fall time of generated pulse.

**Figure 21:** Stable transition times of the generated pulse. The fast speed is due to the use of direct port manipulation rather than digitalWrite.

### 6.2.3 True Delay Between Notes

Another key component of our code design was the use of asynchronous timers to ensure that the pulses corresponding to any note will be perfectly spaced according to that note’s frequency.

Figure 23 displays the measured period of the highest-frequency note (MIDI Note 127), which corresponds to $12543.854$Hz, or an ideal period of $79.720\mu$s. Even though measuring the period with the oscilloscope’s cursors introduces some error in resolution, the apparent period generated by the Arduino in response to MIDI note 127 is $79.6\mu$s, a near perfect replica of the desired result (0 by 0.1% or less).

Figure 24 shows the measured period of the lowest-frequency note (MIDI Note 0), which corresponds to $8.176$Hz or an ideal period of $122.312$ms. At such a large time scale, the oscilloscope only gives resolution to the millisecond. Nevertheless, we were right on the money, measuring a period of $122$ms. For this trial, we increased the pulse width to $1000\mu$s so that the pulses would be visible even at such a zoomed-out view. Nevertheless, this increased pulse width did not affect the overall period
of the pulses (as one might expect by the delay call in the blinkLED function), further proof that the asynchronicity is working as desired.

6.2.4 Support for Multiple Notes

The primary reason for using asynchronous timers to control pulse generation was to support multiple notes at once. Figure 25 demonstrates two concurrent notes. Visually, it is hard to distinguish more than two concurrent notes; however, this system supports up to nine concurrent notes, each running on a different one of the Arduino Due’s nine asynchronous timers.

![Figure 22: Instantaneous transition times of the generated pulse. The fast speed is due to the use of direct port manipulation rather than digitalWrite.](image)

6.2.5 Transmission Across Fiber Optics

While the Arduino was able to generate these output pulses, we needed a way to quickly send them over to the logic circuit. We used a fiber optic transmission system to send these signals from the Arduino’s output to the logic circuit. While the original design featured a transistor (to raise the 3.3V output from the Arduino to the more conventional 5V input to the fiber optic transmitter), we ultimately decided to power the fiber optic transmitter directly from the Arduino’s output pin to save time and remove complexity.

Figure 26 shows the turn-on and turn-off delay times between the voltage across the fiber optic transmitter (shown in bright white) and
the voltage out of the fiber optic receiver (shown in gray). Interestingly, these components seemed to be switching much slower than their specifications. The fiber optic transmitter specifies a rise and fall time of 100\textit{ns} [35] and the receiver specifies a rise and fall time of 70\textit{ns}. [36]

### 6.3 Primary Circuit

The primary circuit was not successful as a whole. However, portions of the circuit functioned exactly as expected. Unfortunately, data is not readily available for all of the following claims, due to time constraints and a make-it-work-first-record-data-later mindset.

On the power side, the doubling rectifier was very successful. The diodes rectified the AC from the wall socket (although we used a variac as a proxy), and the capacitors smoothed out the bumps in the AC. In fact, the voltage difference across the power supplies looked smoother than the DC input we were using from a power supply to power the logic side.

**Figure 23:** The measured time interval between pulses corresponding to MIDI Note 127. The expected value is 79.72Hz, and the measured value is 79.6Hz—almost perfect. Note: The pulse width is 10\textmu s in this image.
Figure 24: The measured time interval between pulses corresponding to MIDI Note 127. The expected value is 79.72Hz, and the measured value is 79.6Hz—almost perfect. Note: For visualization purposes only, the pulse width is 1000μs in this image.
**Figure 25:** Two concurrent notes being played. The top sequence of notes (first, third, and sixth spike) corresponds to one frequency, and the middle sequence of notes (second, fourth, and fifth spike) corresponds to another.
The current transformer, too, seemed to be functioning well. However, we were not able to induce oscillations in the primary circuit, so it is unclear whether the current transformer would have worked. It did, however, give us an accurate readout of the current rise in an RC circuit (which is how the primary circuit behaved with malfunctioning transistors).

The transistors were able to switch very, very quickly. However, towards the end, we ran into trouble making sure the transistors still functioned as NPNs, because we had perhaps blown them out.
On the logic side, the myriad of protective capacitors and resistors would only be relevant in the case of a ground surge, which we were unable to replicate. All power indicator LEDs and fans worked as expected. The voltage regulators output a constant voltage (after many hours of fiddling and burning ourselves, however). While we were unable to take conclusive data on the output of the flip-flop or the gate drivers, we did get to see the not gate in action.

The output of the current transformer should have been AC, so we approximated it with the output of a variac. To be fair, the current transformer’s output would have been at the frequency of the Tesla coil (in the high KHz to MHz range), whereas the frequency of the variac is 60Hz. So while the variac did not let us test the speed of our logic side, it did allow us to see the progression of logic signals in the logic side.

7 Future Work

I would like to expand on this project in a few ways.

MIDI files are pure binary, which means that they are relatively small. It would be great to load a playlist of MIDI songs onto an SD card and allow the Arduino to read a data stream from the SD card. A switch could control the mode (incoming MIDI stream or built-in MIDI songs). Furthermore, it would be convenient to prototype a circuit that allows for fast switching between MIDI from the piano and MIDI from the computer. As of now, I have to unplug one connection’s MIDI cable and re-plug another’s. However, it would be possible to implement a radio button interface (i.e. only one active at a time) that could allow easy toggling between various MIDI input ports.

While hard, it is theoretically possible to interlace the PWM outputs of two Arduinos, each controlled by an independent MIDI source, allowing for multiple-instrument music played on one Tesla coil. The challenge would be handling overlapping signals. More simply, multiple singing Tesla coils could be constructed, one for each of a band’s common instruments.
8 Acknowledgements

I would like to acknowledge a few people for their contributions to this project. First, I must thank Dr. Dann for his ongoing help and willingness to spend time with me, both in the theoretical and practical realms, analyzing and debugging circuits and ICs. Mr. Ward always demonstrated an interest in the project and offered great advice. Early in the process, we spent a lot of time discussing possible approaches to musically modulating the primary signal. Kayla and the rest of the staff at oneTesla helped me conceptually understand the primary circuit. My parents have been extremely supportive throughout these five months. Lastly, I'd like to thank TI for shipping so many free ‘sample’ ICs.

Works Cited


[26] A Different Kind of Internet and TV. GoogleFiber. https://fiber.google.com/about/


The Menlo Roundtable
Katharine Hanson, Editor
Dr. Charles Hanson, Editor emeritus

For the online version of The Roundtable, go to roundtable.menloschool.org.
Russian America: Canvassing the State Ticket
Political Cartoon, Unknown (circa 1867)
Granger Collection (granger.com)