Lexer code

```
def lex(source_code):
    lexemes = []
    i = 0
    while i < len(source_code):
        # returns next non-whitespace index
        i = skip_whitespace(source_code, i)
        if source_code[i] in constants:
            lexemes.append(source_code[i])
        elif source_code[i].isdigit():
            num = source_code[i]
            i += 1
            while i < len(source_code) and source_code[i].isdigit():
                num += source_code[i]
                i += 1
            lexemes.append(num)
            i -= 1 # undo last increment because now not pointing to integer
        else:
            raise SyntaxError("Unknown character: {}".format(source_code[i]))
        i += 1
    return lexemes
```

```
def skip_whitespace(s, idx):
    while s[idx] in ' \t\r\n':
        idx += 1
    return idx
```

```
LPAREN = r"("
RPAREN = r")"
OPPLUS = r"+"
OPMINUS = r"-"
OPMUL = r"*"
OPDIV = r"/"

constants = [
    LPAREN, RPAREN,
    OPPLUS, OPMINUS,
    OPMUL, OPDIV
]
```

Lexer - sample outputs

- (1+2) => ['(', '1', '+', '2', ')']
- 10 / 2 => ['10', '/', '2']
- -(1 + -2) => ['-', '(', '1', '+', '-', '2', ')']
- Note: doesn't need to be valid at this step:
  - 1 + (2 + 3 => ['1', '+', '(', '2', '+', '3']
- But does need to have only lexable characters:
  - abc + 1 => Encountered lexing error: SyntaxError('Unknown character: a')